UNIVERSITÉ DE MONTRÉAL

AUTOMATIC AXON AND MYELIN SEGMENTATION OF MICROSCOPY IMAGES AND MORPHOMETRICS EXTRACTION

ALDO ZAIMI

INSTITUT DE GÉNIE BIOMÉDICAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE BIOMÉDICAL)

AVRIL 2018

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

AUTOMATIC AXON AND MYELIN SEGMENTATION OF MICROSCOPY IMAGES AND MORPHOMETRICS EXTRACTION

présenté par : ZAIMI Aldo

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. STIKOV Nikola, Ph. D., président

M. COHEN-ADAD Julien, Ph. D., membre et directeur de recherche

Mme CHERIET Farida, Ph. D., membre

DEDICATION

To my family.

"You have to believe in the long term plan you have but you need the short term goals to motivate and inspire you." Roger Federer.

ACKNOWLEDGEMENTS

First, I would like to express my sincere gratitude to my research advisor Julien for giving me the opportunity to work in his lab. His patience, knowledge of the neuroscience field and strong work ethic were greatly appreciated and a source of motivation throughout this project. What started out as a standard image processing project quickly turned into a large collaborative work involving software management, computer vision and deep learning. I had so much to learn during this project, and Julien pushed me to give my best in every aspect of it.

I would like to acknowledge the *Natural Sciences and Engineering Research Council of Canada* (NSERC) who financially supported me during the last year.

I would also like to thank the members of my thesis committee, Prof. Farida Cheriet and Prof. Nikola Stikov, who kindly accepted to read and evaluate my work. Your insightful comments are greatly appreciated.

This project would have not been possible without the continuous help and support from my fellow colleagues. I am very grateful and humbled to have met such amazing people from various backgrounds, expertise and cultures. The following people were more closely involved in my project and deserve a special mention: Ariane Saliani, who did all the tissue preparation and microscopy imaging; Tanguy Duval, who was strongly invested in the development of *AxonSeg*; Christian Perone, who helped me in the deep learning and software engineering parts of my project; Victor Herman, Pierre-Louis Antonsanti and Maxime Wabartha, who worked as interns in the lab and did some tremendous work on the *AxonDeepSeg* software.

I also want to thank Nikola for his implication, support and for his insightful feedback throughout this project. The microscopy data he shared contributed to the success of this project.

Finally, I want to thank all the other members of the lab: Agah, Alexandru, Alexandros, Atef, Benjamin, Charley, Darya, Dominique, Francisco, Gabriel, George, Harris, Nibardo, Ryan, Sara, Simon, Stephanie, Tommy and Tung.

RÉSUMÉ

Dans le système nerveux, la transmission des signaux électriques se fait par l'intermédiaire des axones de la matière blanche. La plupart de ces axones, aussi connus sous le nom de fibres nerveuses, sont entourés par la gaine de myéline. Le rôle principal de la gaine de myéline est d'accroître la vitesse de transmission du signal nerveux le long de l'axone, un élément crucial pour la communication sur de longues distances. Lors de pathologies démyélinisantes comme la sclérose en plaques, la gaine de myéline des axones du système nerveux central est attaquée par des cellules du système immunitaire. Ceci peut conduire à la dégénérescence de la myéline, qui peut se manifester de diverses façons : une perte du contenu en myéline, une diminution du nombre d'axones myélinisés ou même des dommages axonaux.

La microscopie à haute résolution des tissus myélinisés offre l'avantage de pouvoir imager la microstructure du tissu au niveau cellulaire. L'extraction d'information quantitative sur la morphologie passe par la segmentation des axones et gaines de myélines composant le tissu sur les images microscopiques acquises. L'extraction de métriques morphologiques des fibres nerveuses à partir d'image microscopiques pourrait contribuer à plusieurs applications intéressantes : documentation de la morphométrie sur différentes espèces et tissus, étude des origines et effets des maladies démyélinisantes, et validation de nouveaux biomarqueurs d'Imagerie par Résonance Magnétique sensibles au contenu en myéline dans le tissu.

L'objectif principal de ce projet de recherche est de *concevoir, implémenter et valider un framework de segmentation automatique d'axones et de gaines de myéline sur des images microscopiques et d'en extraire des morphométriques pertinentes*. Plusieurs approches de segmentation ont été explorées dans la littérature, mais la plupart ne sont pas totalement automatiques, sont conçues pour une modalité de microscopie spécifique, ou bien leur implémentation n'est pas publiquement disponible pour la communauté scientifique. Deux *frameworks* de segmentation ont été développés dans le cadre de ce projet : *AxonSeg* et *AxonDeepSeg*.

Le *framework AxonSeg* (https://github.com/neuropoly/axonseg) se base sur une approche de traitement d'image classique pour la segmentation. Le pipeline de segmentation inclut une transformée de type *extended-minima*, un modèle d'analyse discriminante combinant des *features* de forme et d'intensité, un algorithme de détection de contours et un double algorithme de

contours actifs. Le résultat de la segmentation est utilisé pour l'extraction de morphométriques. La validation du *framework* a été réalisée sur des échantillons de microscopie optique, microscopie électronique et microscopie Raman stimulée (CARS).

Le *framework AxonDeepSeg* (https://github.com/neuropoly/axondeepseg) utilise plutôt une approche basée sur des réseaux neuronaux convolutifs. Un réseau convolutif a été conçu pour la segmentation sémantique des axones myélinisés. Un modèle de microscopie électronique à balayage (MEB) a été entraîné sur des échantillons de moelle épinière de rat et un modèle de microscopie électronique à transmission (MET) a été entraîné sur des échantillons de corps calleux de souris. Les deux modèles ont démontré une haute précision pixel par pixel sur les échantillons test (85% sur le MEB de rat, 81% sur le MEB d'humain, 95% sur le MET de souris, 84% sur le MET de macaque). On démontre également que les modèles entrainés sont robustes aux ajouts de bruit, au flou et aux changements d'intensité. Le modèle MEB de *AxonDeepSeg* a été utilisé pour segmenter une coupe transversale complète de moelle épinière de rat et les morphométriques extraites à partir des tracts de la matière blanche correspondaient bien aux tendances rapportées dans la littérature. *AxonDeepSeg* a démontré une plus grande précision au niveau de la segmentation lorsque comparé à *AxonSeg*. Les deux outils logiciels développés sont *open source* (licence MIT) et donc à disposition de la communauté scientifique.

Des futures itérations sont prévues afin d'améliorer et d'étendre ce travail. Les objectifs à court terme sont l'entraînement de nouveaux modèles pour d'autres modalités de microscopie, l'entraînement sur des *datasets* plus larges afin d'améliorer la généralisation et la robustesse des modèles, et l'exploration de nouvelles architectures de réseaux neuronaux. De plus, les modèles de segmentations développés jusqu'à maintenant ont seulement été testés sur des images de tissus sains. Un développement futur important serait de tester la performance de ces modèles sur des échantillons démyélinisés.

ABSTRACT

In the nervous system, the transmission of electrical signals is ensured by the axons of the white matter. A large portion of these axons, also known as nerve fibers, is surrounded by a myelin sheath. The main role of the myelin sheath is to increase the transmission speed along the axons, which is crucial for long distance communication. In demyelinating diseases such as multiple sclerosis, the myelin sheath of the central nervous system is attacked by cells of the immune system. Myelin degeneration caused by such disorders can manifest itself in different ways at the microstructural level: loss of myelin content, decrease in the number of myelinated axons, or even axonal damage.

High resolution microscopy of myelinated tissues can provide in-depth microstructural information about the tissue under study. Segmentation of the axon and myelin content of a microscopy image is a necessary step in order to extract quantitative morphological information from the tissue. Being able to extract morphometrics from the tissue would benefit several applications: document nerve morphometry across species or tissues, get a better understanding of the origins of demyelinating diseases, and validate novel magnetic resonance imaging biomarkers sensitive to myelin content.

The main objective of this research project is to *design, implement and validate an automatic axon and myelin segmentation framework for microscopy images and use it to extract relevant morphological metrics*. Several segmentation approaches exist in the literature for similar applications, but most of them are not fully automatic, are designed to work on a specific microscopy modality and/or are not made available to the research community. Two segmentation frameworks were developed as part of this project: *AxonSeg* and *AxonDeepSeg*.

The *AxonSeg* package (https://github.com/neuropoly/axonseg) uses a segmentation approach based on standard image processing. The segmentation pipeline includes an extended-minima transform, a discriminant analysis model based on shape and intensity features, an edge detection algorithm, and a double active contours step. The segmentation output is used to compute morphological metrics. Validation of the framework was performed on optical, electron and CARS microscopy.

The *AxonDeepSeg* package (https://github.com/neuropoly/axondeepseg) uses a segmentation approach based on convolutional neural networks. A fully convolutional network architecture was designed for the semantic 3-class segmentation of myelinated axons. A scanning electron microscopy (SEM) model trained on rat spinal cord samples and a transmission electron microscopy (TEM) model trained on mice corpus callosum samples are presented. Both models presented high pixel-wise accuracy on test datasets (85% on rat SEM, 81% on human SEM, 95% on mice TEM and 84% on macaque TEM). We show that *AxonDeepSeg* models are robust to noise, blurring and intensity changes. *AxonDeepSeg* was used to segment a full rat spinal cord slice, and morphological metrics extracted from white matter tracks correlated well with the literature. The *AxonDeepSeg* framework presented a higher segmentation accuracy when compared to *AxonSeg*. Both *AxonSeg* and *AxonDeepSeg* are open source (MIT license) and thus freely available for use by the research community.

Future iterations are planned to improve and extend this work. Training of new models for other microscopy modalities, training on larger datasets to improve generalization and robustness, and exploration of novel deep learning architectures are some of the short-term objectives. Moreover, the current segmentation models have only been tested on healthy tissues. Another important short-term objective would be to assess the performance of these models on demyelinated samples.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AVF	Axon Volume Fraction
CARS	Coherent Anti-Stokes Raman Scattering
CNN	Convolutional Neural Network
CNS	Central Nervous System
FCN	Fully Convolutional Network
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
MVF	Myelin Volume Fraction
OM	Optical Microscopy
PNS	Peripheral Nervous System
SEM	Scanning Electron Microscopy
TEM	Transmission Electron Microscopy

CHAPTER 1 INTRODUCTION

The central nervous system (CNS) consists of the brain and the spinal cord and is the integration and coordination center of the body. Transmission of the information between the different components of the CNS is done via the axons. These axons are usually surrounded by a myelin sheath. The myelin sheath is essential to the transmission of nerve impulses: it significantly increases the transmission speed by which the electrical signals propagate. Several CNS disorders can affect the signal propagation of the nerve fibers. Demyelinating diseases such as multiple sclerosis (MS) attack the myelin sheaths of the white matter axons, causing important sensory and motor deficiencies. MS has a higher rate in Canada when compared to other countries: 1 in 340 Canadians are living with the disease [1]. It is typically diagnosed in young adults and requires monitoring of the nature and severity of the symptoms. Therefore, the development of robust diagnosis and monitoring methods is key to the clinical outcome. Magnetic Resonance Imaging (MRI) is the most commonly used clinical tool for the diagnosis and monitoring of MS. Macrostructural MS lesions can be localized in typical MRI contrasts such as T1-weighted or T2-weighted.

At the microstructural level, demyelination can cause loss of myelin content and lead to axonal damage and axonal loss. High resolution microscopy is capable of imaging the individual axons and myelin sheaths of a tissue. Having access to such detailed information enables extraction of morphological measurements such as the axon diameter, the myelin thickness and the axon density. However, to obtain that microstructural information, robust segmentation of each axon and myelin sheath in the image is needed. Manual segmentation is not a reliable option, as a spinal cord cross section can contain millions of nerve fibers. Therefore, having robust and automatic algorithms that can perform the segmentation task is essential. The extraction of morphometric measurements from myelinated axons can contribute to many applications: documentation of the nerve morphometry across species and tissues, study of demyelination in animal models, or validation of novel MRI techniques that aim at quantifying the axonal and/or myelin content of a tissue.

The **main objective** of this research project is to *develop a framework that performs automatic axon and myelin segmentation from microscopy data*. The framework should be validated on the commonly used microscopy modalities (scanning electron microscopy,

transmission electron microscopy, optical microscopy), tissues of the CNS (spinal cord and brain white matter), and species (e.g. rat, mouse, cat, human). Moreover, the framework should be capable of extracting morphometrics from the segmentation output. Another underlying goal of this project is to make the resulting implementations open access. Most of the existing frameworks designed for similar applications are not available to the research community.

1.1 Organization of the thesis

The present thesis is organized as follows. Chapter 2 presents a critical literature review, including anatomy of nerve fiber cells of the CNS, pathologies affecting the CNS, histology and microscopy of CNS tissues, existing segmentation approaches for similar applications, axon and myelin morphometrics, and motivations of the morphometrics extraction. The general and specific research objectives are also identified. Chapter 3 gives an overview of the methodology used in this project. Chapters 4 and 5 present the two articles published during this project:

1. **Zaimi, A.**, Duval, T., Gasecka, A., Côté, D., Stikov, N., & Cohen-Adad, J. (2016). AxonSeg: Open Source Software for Axon and Myelin Segmentation and Morphometric Analysis. *Frontiers in Neuroinformatics*, *10*, 37. https://doi.org/10.3389/fninf.2016.00037
2. **Zaimi, A.**, Wabartha, M., Herman, V., Antonsanti, P.-L., Perone, C. S., & Cohen-Adad, J. (2018). AxonDeepSeg: automatic axon and myelin segmentation from microscopy data using convolutional neural networks. *Scientific Reports*, *8*(1), 3816. https://doi.org/10.1038/s41598-018-22181-4

Chapter 6 presents supplementary methods and results that were done after the submission of the *AxonDeepSeg* paper. Chapter 7 presents a general discussion of the project. Chapter 8 concludes by summarizing the method and results and lists possible future developments of the project.

CHAPTER 2 LITERATURE REVIEW

The following chapter presents a literature review of the project. Section 2.1 presents the anatomy and physiology of the central nervous system (CNS). Section 2.2 gives a brief overview of the disorders affecting the CNS. Section 2.3 presents the histology and microscopy imaging of myelinated tissues of the CNS. Section 2.4 presents an overview of the existing standard image processing and deep learning approaches for the segmentation of myelinated fibers from microscopy images. Limitations of previous work are identified. Section 2.5 presents the main morphological metrics of interest that can be obtained from the segmentation of myelinated fibers. Section 2.6 identifies and discusses three major motivations (i.e. applications) of morphometrics extraction. Finally, section 2.7 defines the general and specific objectives of this project.

2.1 Anatomy and physiology of the central nervous system

2.1.1 The central nervous system

The vertebrate central nervous system (CNS) consists of two major structures: the brain and the spinal cord (see **Figure 2.1**). The brain controls and commands the body functions, while the spinal cord, among other functions, ensures the propagation of information between the brain and the peripheral nervous system (PNS) nerves. For instance, motor commands originating from the brain travel the spinal cord to reach the muscles, while the sensory information captured by tissues such as the skin travels the spinal cord and then reaches the brain to be processed.

2.1.2 The structure and function of the neuron

The neuron, also known as nerve cell, is the primary unit of the nervous system. It can receive, process and transmit information in the form of electrical signals. A typical neuron consists of three main parts: the cell body (also called soma), the dendrites and the axon (also called nerve fiber). The cell body contains the nucleus along with other organelles such as the endoplasmic reticulum, the Golgi apparatus and the mitochondria. In most cases, neurons receive signals via the dendrites and pass them through the axon.

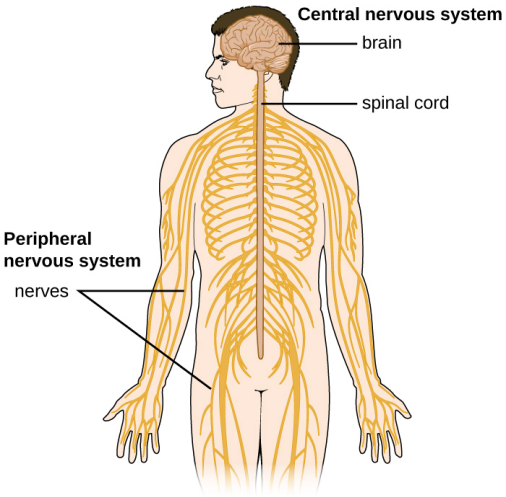

Figure 2.1 : Anatomy of the nervous system [2]. The central nervous system (CNS) consists of the brain and the spinal cord.

Axons may be surrounded by a myelin sheath that acts as an electrical insulator. These axons wrapped by a myelin sheath are called myelinated axons. In the CNS, myelinated axons can be found in the white matter. The myelin sheath is mostly composed of lipids (between 70-85%) and proteins (between 15-30%) [3]. The main role of the myelin sheath is to increase the transmission speed of the signal propagation along the nerve fiber. The transmission relies on saltatory conduction, a process by which action potentials jump from one node of Ranvier (portions where the axon is left uncovered by myelin) to the next (see **Figure 2.2**) [4]. In the CNS, the myelin sheath is produced by the Schwann cells, while in the PNS it is produced by oligodendrocytes [5]. In a cross section of myelinated tissue (i.e. perpendicular to the axon orientation), the myelin sheaths appear as concentric layers around the axon (see **Figure 2.3**) [6].

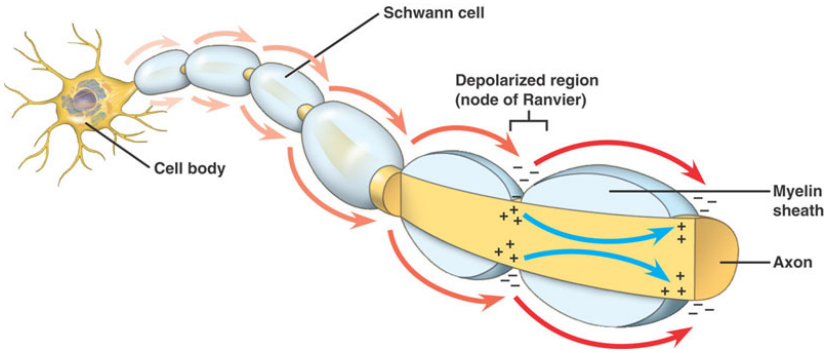

Figure 2.2: Illustration of a neuron with a myelinated axon [7]. The electrical signal is transmitted by saltatory conduction.

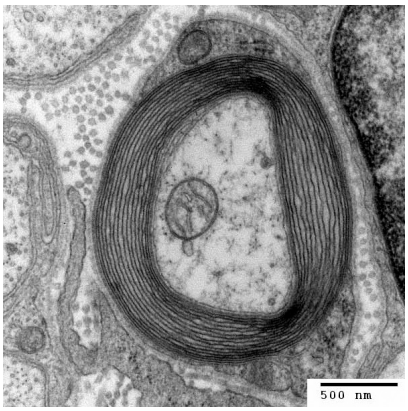

Figure 2.3: Cross section of a nerve fiber obtained by transmission electron microscopy (TEM) [8]. The myelin sheaths appear as thin concentric layers wrapped around the axon.

2.1.3 Myelinated axons in the brain

The brain is composed of two main tissues: the gray matter and the white matter. The white matter is made of bundles of myelinated axons and is located beneath the gray matter cortex (see **Figure 2.4 (a)**) [9]. The nerve fibers of the white matter connect the different areas of the gray matter. The corpus callosum (see **Figure 2.4 (b)**) is the largest white matter structure of the brain, containing more than 200 millions nerve fibers (myelinated and unmyelinated) [10]. Its role is to interconnect the left and right hemispheres of the brain.

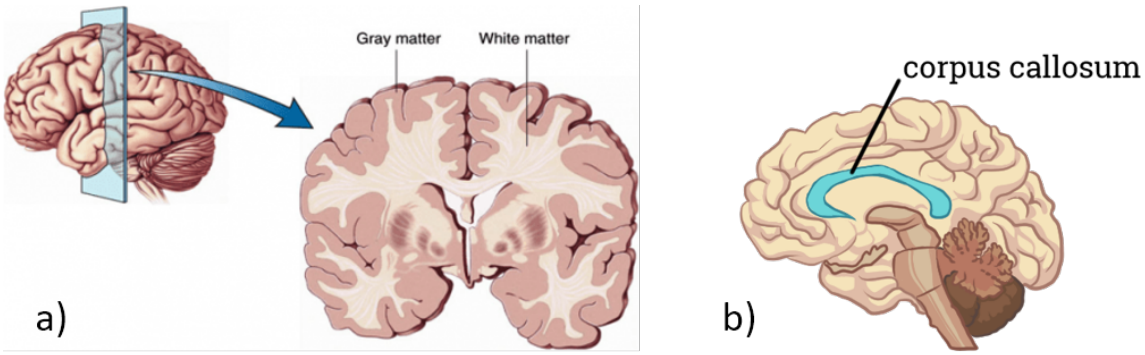

Figure 2.4: Illustration of the two main tissues of the brain: gray and white matter (a) [11], and location of the corpus callosum in the brain (b) [12].

2.1.4 Myelinated axons in the spinal cord

In the spinal cord, the gray matter is located at the center and is surrounded by the white matter (opposite to the brain). In a cross section of the spinal cord, the gray matter appears as a butterfly-shaped structure. The spinal cord can be divided into regions and segments. For instance, in the human spinal cord there are 4 regions divided into segments: 8 cervical segments, 12 thoracic segments, 5 lumbar segments and 5 sacral segments. The shape and relative area of the white and gray matter varies along the spinal cord levels (see **Figure 2.5**).

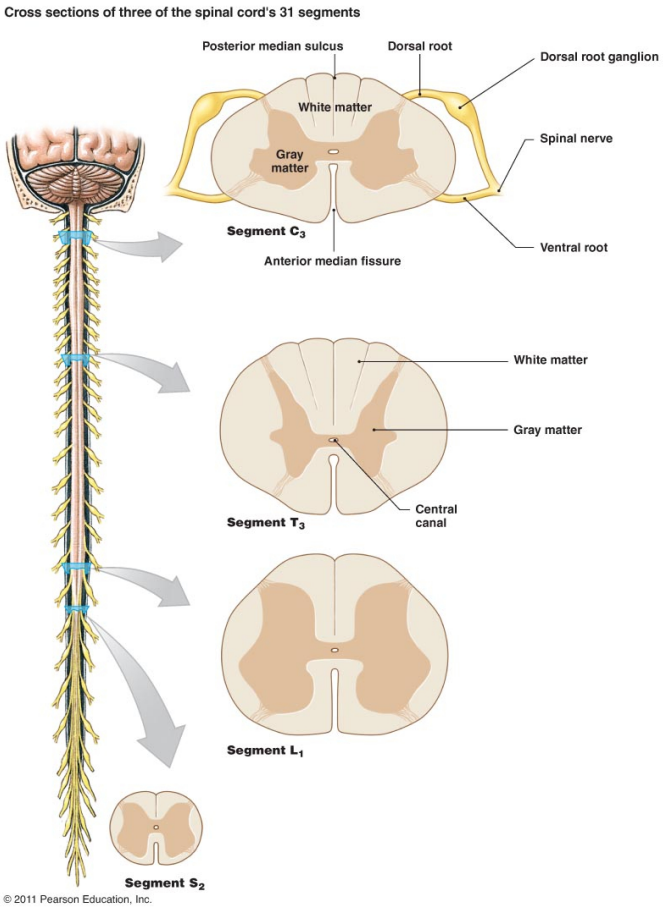

Figure 2.5: Diagram illustrating the variability in shape and size of the gray and white matter across the different spinal cord segments [13].

The white matter is the main source of myelinated axons in the spinal cord. It can be divided into three columns based on the location: ventral, dorsal and lateral columns. It can also be divided into tracts, which consist of bundles of axons that have a specific function. Ascending tracts deliver sensory information to the brain, while descending tracts transmit motor

information to the periphery (see **Figure 2.6**). For instance, the gracile fasciculus tract transmits limb and trunk position sensations, deep touch, visceral pain and vibration located below vertebral level T6 [14].

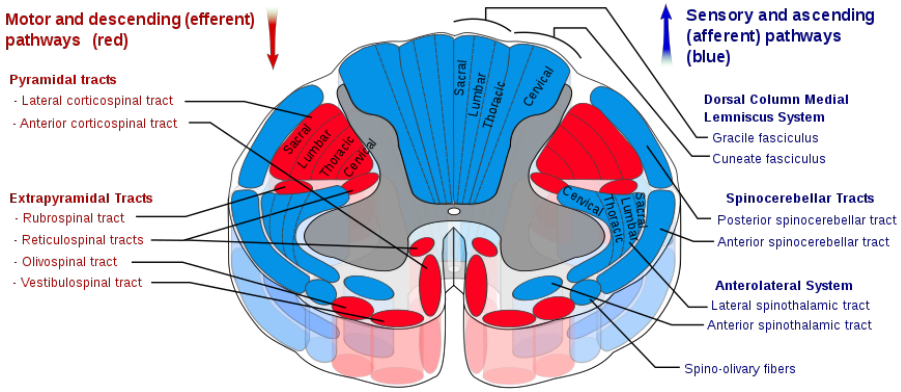

Figure 2.6: Ascending (blue) and descending (red) white matter tracts of the human spinal cord [15].

2.2 Central nervous system diseases

The central nervous system is the target of many disorders such as traumatic injuries, infections, tumors, degenerative diseases and demyelinating diseases. Examples of degenerative diseases include Alzheimer's disease, amyotrophic lateral sclerosis (ALS) and Parkinson's disease. The most prominent example of demyelinating disease is multiple sclerosis (MS).

2.2.1 Multiple sclerosis

Multiple sclerosis is a chronic inflammatory disease that attacks the myelin sheath of myelinated axons in the CNS, a process known as demyelination [16]. Although the ultimate cause of MS is not clearly known yet, it is believed that MS lesions are initiated by autoreactive T lymphocytes of the immune system that attack myelinated nerve fibers (see **Figure 2.7**) [17]. The destruction of the myelin sheath affects signal transmission and leads to important sensory and motor deficits [18]. In histopathological studies, demyelination is usually characterized by lower myelin content of the tissue (i.e. lower thickness of the myelin sheaths) (see **Figure 2.8**).

Magnetic Resonance Imaging (MRI) is a commonly used clinical tool for the diagnosis and monitoring of MS [19]. Due to the difference in composition between healthy and demyelinated tissues, MS lesions can be detected with good sensitivity by conventional MRI

[20]. For instance, MS lesions imaged by T2-weighted MRI appear as regions with high signal [21].

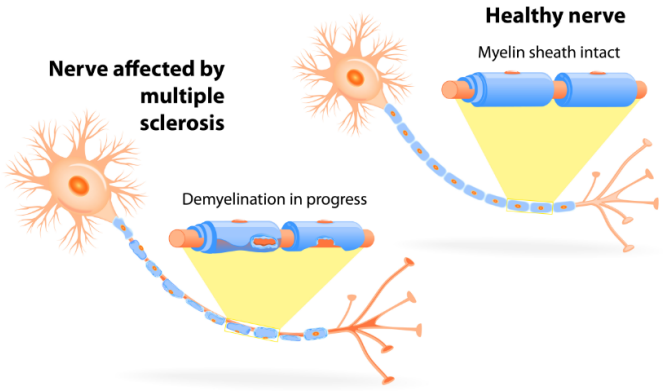

Figure 2.7: Schematic illustration of demyelination in myelinated nerve fibers [22].

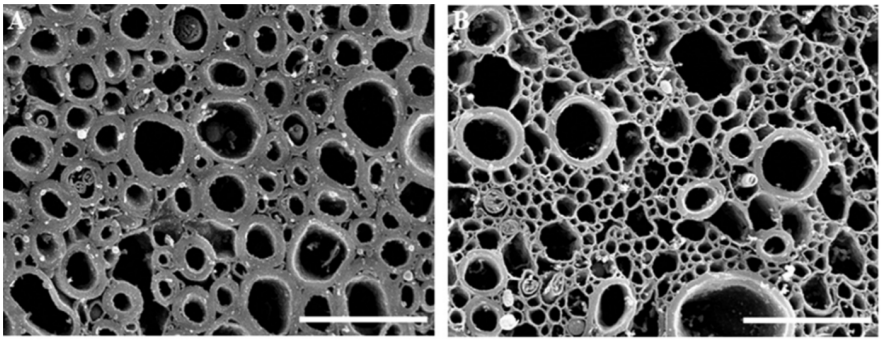

Figure 2.8: Cuprizone-induced demyelination in the corpus callosum of a mouse brain, imaged by scanning electron microscopy (SEM). Left: control. Right: treatment with cuprizone for 4 weeks. Scale bars: 2 µm. Adapted from [23].

2.3 Histology and microscopy of the central nervous system

While MRI is used for clinical diagnosis and monitoring of CNS diseases, histology is the preferred method for observation of microstructural properties and changes in the tissue *ex vivo* [24]. Histological preparations can be imaged by microscopy at very high resolution to visualize the individual axons and myelin sheaths of a myelinated tissue.

2.3.1 Tissue preparation for microscopy

Several histological protocols for the preparation of nervous tissue exist in the literature. The tissue preparation is a critical step as it can be affected by many factors (e.g. fixation method,

duration of fixation, staining) [25]. The standard procedure is to fix the tissue (either by immersion or perfusion) using a mixture of glutaraldehyde (Ga) and paraformaldehyde (PFA). Osmification is then typically done by using osmium tetroxide (OsO_4) to fix the lipids (i.e. myelin sheaths of myelinated tissues) and stain the tissue for subsequent microscopy imaging [25]. The osmified sample is then dehydrated and embedded in epoxy resin for microscopy imaging.

2.3.2 Microscopy modalities

Several microscopy modalities have been employed for microstructure observations of white matter tissues. Most microstructural studies of the white matter have been based on optical (light) and electron microscopy, but novel imaging techniques are being investigated and validated.

2.3.2.1 Optical microscopy

Standard optical (light) microscopy (OM) of nerve fibers has been a popular and accessible technique in earlier microstructure studies, especially when combined with toluidine blue staining [26-27]. It is well suited to large scale histology (e.g. imaging of whole spinal cord slice) [28], but its limited resolution (0.2 μm [29]) does not allow imaging of the smallest nerve fibers.

2.3.2.2 Transmission electron microscopy

Transmission electron microscopy (TEM) is often considered as a gold-standard imaging technique for the observation of nerve fiber microstructure [30] due to its very high resolution limit (around 0.2 nm [31]). It is the most widely used microscopy modality to study white matter microstructure. However, it has a small field of view and tissue preparation requires very thin sample slices.

2.3.2.3 Scanning electron microscopy

Scanning electron microscopy (SEM) is another modality based on the interaction between the tissue and an electron beam. Its resolution limit is around 10 nm [31]. Unlike TEM, SEM is suited to large scale histology as it allows larger fields of view during acquisition. For

instance, a whole rat spinal cord slice can be imaged by imaging regions of around 500×500 μm^2. Imaged sub-sections can then be stitched together to reconstruct the whole slice.

2.3.2.4 Coherent anti-Stokes Raman scattering microscopy

Coherent anti-Stokes Raman scattering (CARS) microscopy is a recent technique that uses the vibrational contrast of the molecules of the different tissue structures [32]. The main advantage of CARS microscopy is its chemical selectivity and high sensitivity to the myelin content of the tissue [33]. Moreover, it does not require labeling agents (i.e. dye-free), which also reduces the risk of artifacts due to the tissue fixation and/or staining [34]. Its resolution is comparable to standard light microscopy (0.28 μm [35]). Due to its sensitivity to myelin, it has been used to assess myelin degradation in demyelination models [36].

2.3.2.5 Appearance of myelinated axons in microscopy images

Figure 2.9 gives an overview of the appearance of myelinated nerve fibers in the most common microscopy modalities. In OM and TEM, myelin sheaths have lower intensity than axons and appear dark, while in SEM and CARS, the contrast is inverted (i.e. axons appear dark). Due to the high resolution of TEM, it is possible to observe internal axonal structures (i.e. organelles). It is also worth noting that in CARS samples, axons and background have the same appearance (due to CARS being only sensitive to myelin content).

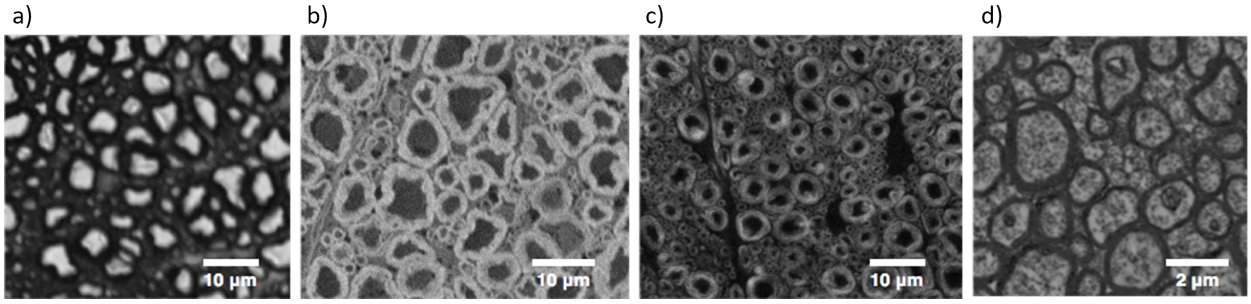

Figure 2.9: Examples of myelinated axons in common microscopy modalities. Samples of cat spinal cord from optical microscopy (a), rat spinal cord from scanning electron microscopy (b), rat spinal cord from CARS microscopy (c) and mouse corpus callosum from transmission electron microscopy (d) are presented.

2.4 Axon and myelin segmentation

The segmentation of axons and myelin sheaths in microscopic images of myelinated tissues is a more specific application of the standard cell segmentation [37]–[39]. It is a multi-class segmentation: the goal is to determine, for each pixel, if it belongs to the axon, myelin or background class. This application offers many interesting challenges: multi-instance segmentation (a small 50×50 μm^2 sample of SEM spinal cord can contain over 100 nerve fibers), touching objects (touching axon-myelin pairs and touching neighboring myelin sheaths), large variability of axons/myelin in shape and size, presence of other structures in the extracellular space (i.e. cellular debris, vessels), presence of organelles in the intracellular space in very high resolution microscopy (i.e. TEM), tissue preparation artifacts (e.g. myelin sheath disruptions and nerve fiber shrinkage) and imaging artifacts (e.g. illumination inhomogeneities).

Most past work has focused on segmentation techniques based on standard image processing such as histogram-based methods (thresholding), edge detection, region growing, energy-based methods (active contours), partial differential equation-based methods (level sets), graph portioning methods (Markov random fields) and watershed. Section 2.4.1 presents an overview of the most relevant past work on axon and myelin segmentation from microscopy images.

In recent years, deep learning-based methods have gained huge popularity, especially when it comes to computer vision tasks. Applications based on the training of convolutional neural networks (CNNs) have been outperforming many traditional algorithms in tasks such as image recognition [40] and semantic segmentation [41]. Section 2.4.2 presents a quick introduction to CNNs, and section 2.4.3 presents example applications of deep learning to the axon and myelin segmentation challenge.

2.4.1 Standard image processing approaches

Earlier experiments on axon and myelin segmentation were based on simple algorithms such as thresholding to obtain a binary mask of the myelin sheaths [27], [42], [43]. Romero and collaborators [27] first use a local histogram thresholding algorithm to separate the myelin rings from the rest of the image. Then, the binary image is processed by morphological operators and

filtered by using size and circularity criteria. Finally, touching objects are separated by using an Euclidian distance transform.

More and collaborators [44] start by denoising the image using a combination of morphological operators and morphological reconstructions. Their denoising method offers the advantage of preserving edges in the image, unlike standard denoising techniques that rely on convolutional filters. Then, a threshold is applied to obtain a binary image of myelin rings. Discrimination using size, circularity and intensity level criteria is performed on the complement of the myelin rings binary image. Manual segmentation is then suggested to correct the output axon and myelin masks (i.e. add missed axons and remove false positives). Similarly, Liu and collaborators [45] propose the Otsu thresholding method to obtain a preliminary axon segmentation. Then, axon candidates are discriminated using size, circularity and solidity criteria. The outer myelin sheath contour is estimated by applying an active contour algorithm starting from the axon contours. In other cases, the active contour algorithm is preceded by an elliptical Hough transform [46].

A simple semi-automatic region growing algorithm proposed by Gierthmuehlen and collaborators [47] consists of assigning seeds to all axons of the sample. A more robust region growing approach has been proposed by Zhao and collaborators [48]. The initial seeds of the algorithm are defined by the intensity of the myelin peak in the image histogram. Once the myelin candidates are determined, axon candidates are detected by region labeling. Size and circularity criteria are used to eliminate false positives. Morphological filtering is used to clean small debris in the axon mask and smooth contours. Finally, adjacent myelin sheaths are separated by computing the maximum average gradient magnitude of the outer myelin sheath boundary.

Begin and collaborators [33] opted for a distinct pipeline for axon and myelin segmentation. Axon segmentation is done by applying an extended-minima algorithm to identify axon candidates. An active contours algorithm is then used to refine the edges of the axon candidates, and discrimination based on shape features (circularity and solidity) is performed. For the myelin segmentation, the outer boundary of the myelin sheaths is detected based on intensity changes (edge detection). Then, discrimination is done based on the area overlap between neighboring myelin objects and touching myelin pairs are separated by a watershed algorithm.

2.4.2 Quick introduction to convolutional neural networks

Machine learning, which is a subfield of artificial intelligence, is a general term that designates algorithms capable of deriving rules from data in order to perform decision tasks such as classification [49]. Machine learning algorithms try to learn useful representations of the data. Deep learning is a subfield of machine learning in which the goal is to learn successive hierarchical layers of representations. The advantage of deep learning over traditional machine learning algorithms is that it automates the feature engineering part of the process. In other words, there is no need to hand-select the relevant features for a given application: the deep learning model will learn these features by itself in a structured and hierarchical way. Deep learning networks rely on three main elements: (i) a parametrization by learnable weights (also called parameters or units of the network); (ii) a loss function to assess the performance of the network and (iii) an optimization algorithm to update the weights depending on the loss function (e.g. backpropagation) [50].

A deep learning framework particularly adapted to computer vision tasks is the convolutional neural network. Most CNN architectures consist of the following building blocks: (i) convolutional layers; (ii) pooling layers and (iii) fully connected layers (see **Figure 2.10**) [51].

Convolutional layers are based on the traditional convolution operation. The goal is to determine filters that will capture the most relevant features of the image for a given task. Depending on the features expected in the image (i.e. learned filters), units of the convolutional layers will fire (i.e. emit a high signal) when they detect such features in an input image patch. These features may be simple edges in the first layer representation, compositions of edges in the second layer representation and more complex textures or structures in the following layers (see **Figure 2.11**).

Activation functions are typically used at the end of the convolutional layers. These activation functions introduce nonlinear transformations into the network, allowing it to learn more complex features by adding some sparsity [52]. These functions decide if a neuron of the network (i.e. unit) should be activated or not (i.e. if the information received by the neuron is relevant and should be kept and propagated into the next layer). The activation function can be seen as a threshold that needs to be reached in order to make the unit fire (the concept of activation functions was in part inspired by the neurons of the nervous system that fire and

propagate the signal when the action potential reaches a certain level). The most commonly used activation function is the rectified linear unit (ReLU). The ReLU function *f(x)* equals 0 when the input is smaller than 0 and equals *x* otherwise (i.e. when the input of a unit is negative, that unit will be converted to 0 and not activated). Other activation functions such as the sigmoid or the tanh can also be used.

Pooling layers aim at downsampling the learned feature representations to obtain more abstract representations that are translation invariant (i.e. better generalization of the feature detection over the entire input image). Various types of pooling layers exist. Max pooling is the most widely used pooling method: it consists of taking the maximum value of a certain region of a matrix (see **Figure 2.12**). Another important concept in the pooling layers is the stride, which defines the step size of the kernel when applying an operation to an input matrix. For instance, a standard convolution operation has a stride of 1. Most pooling layers in CNNs use strides greater than 1 in order to obtain a downsampling of the feature space. For instance, a max pooling operation of stride 2 would map an input matrix space into a downsampled space of factor 2 (see **Figure 2.12** for a concrete example).

Finally, fully connected layers are typically added at the end of the architecture for classification tasks to translate the image representations into an appropriate output. For instance, the output layer of a digit classification network may be a probability score indicating the likelihood of the input image being one of the 10 possible digits. In most multi-class classification problems (i.e. semantic segmentation of images), this last layer can be a softmax activation: for each pixel of the input image, the output is a probability of the pixel being one of the possible classes. Typically, the final classification of the pixel is done by picking the class that yields the highest probability.

The number of convolutional layers (depth), the number of filters (i.e. features learned) per convolutional layer and the kernel size of the filters (e.g. 3×3 vs 5×5) are examples of hyperparameters of the network. Hyperparameters require optimization during the design step of the network architecture depending on the application.

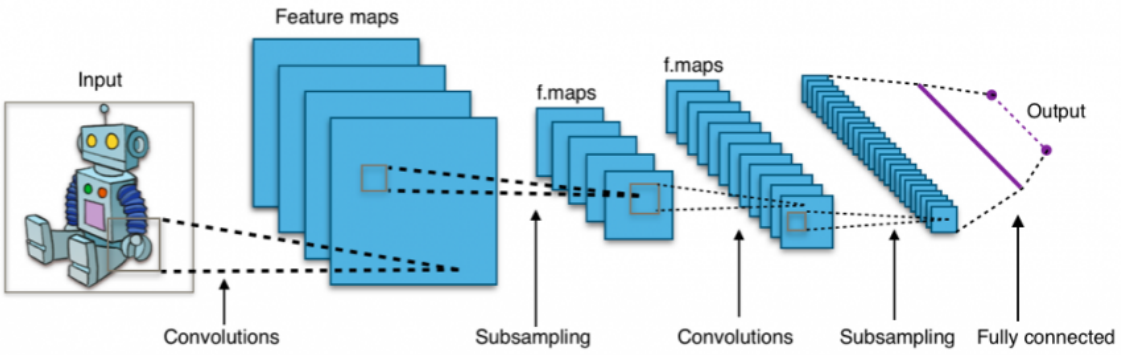

Figure 2.10: Illustration of a typical convolutional neural network architecture used for classification tasks [53]. The input image usually passes through a sequence of convolutional and pooling layers. For classification tasks, a fully connected layer is usually added. The output here can be a vector of probabilities for each object (i.e. class) of the classification task.

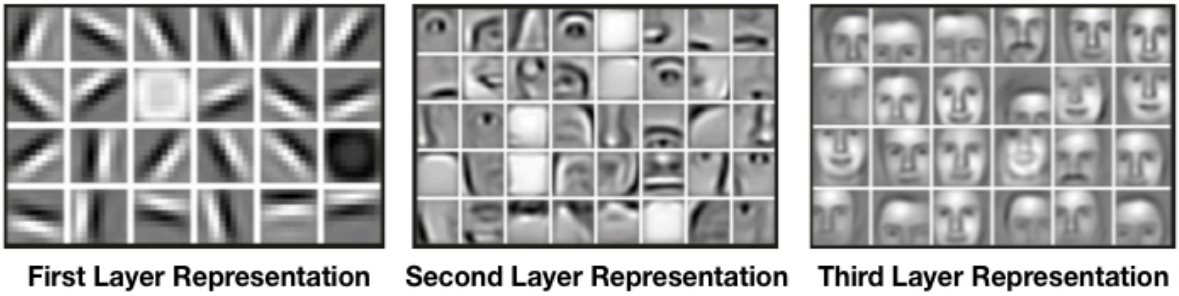

Figure 2.11: Examples of feature layers in a convolutional neural network [54]. The first layers capture simple features such as edges, while deeper layers typically learn more complex and abstract representations of the image (i.e. textures and objects).

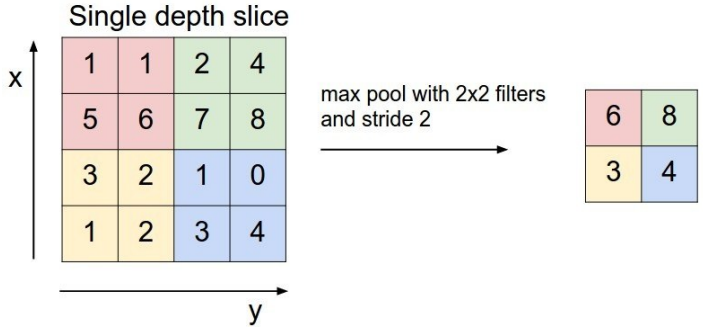

Figure 2.12: Illustration of the max pooling downsampling method [55]. A max pooling of kernel size 2×2 and stride of 2 will convert a 4×4 input matrix to a 2×2 output. Each unit (i.e. element) of the output matrix is the maximum value of the corresponding 2×2 region of the input.

2.4.3 Deep learning based approaches

In the axon/myelin segmentation application, we are interested in performing semantic segmentation (i.e. label each pixel as either axon, myelin or background). It has been demonstrated that convolutional networks can achieve state-of-the-art results in semantic segmentation problems [41]. CNNs have been used for instance for membrane segmentation [56] and cell segmentation [57] in electron microscopy images. The elegant U-Net architecture introduced by Ronneberger and collaborators [57] has gained a lot of popularity for medical segmentation applications. It introduces a way to combine both context and localization to obtain a more robust segmentation of the structures of interest.

Very few applications of CNNs to axon/myelin segmentation exist in the literature. Naito and collaborators [58] first perform a preliminary segmentation of myelin sheaths by using a combination of clustering algorithms. Then, for each myelin sheath candidate detected in the image, a binary classification CNN is used to determine whether it is a myelinated nerve fiber or not. The proposed CNN architecture has 3 convolutional layers, 3 pooling layers and a fully connected layer of 256 units. The model is trained on 64×64 patches obtained from 39 sample images and tested on 10 sample images.

Mesbah and collaborators [59] explored several CNN architectures for axon/myelin segmentation in optical images. They presented a deep encoder-decoder CNN model similar to the U-net that achieves up to 82% pixel-wise accuracy in the 3-class segmentation. Input patches of 32×32 are fed into the network. The fully convolutional network (FCN) consists of a combination of convolutional layers (using 3×3 kernels and outputting feature maps of 8, 64 and 256 channels) and downsampling/upsampling layers.

2.4.4 Summary of axon and myelin segmentation methods

Table 2.1 summarizes the various segmentation methods and algorithms used for axon and myelin segmentation in the literature. For each study, the microscopy modality, species and type of tissue are also reported. Segmentation approaches are divided into 2 groups: frameworks based on standard image processing and frameworks based on deep learning (neural networks).

Table 2.1: Summary of the main axon and myelin segmentation methods in the literature. For each reference, the microscopy modality, the species and the type of tissue are also reported.

Reference	Microscopy modality	Species	Tissue	Segmentation algorithm(s)
Methods based on standard image processing				
Romero 2000 [27]	OM	Cat	Sciatic nerve	Thresholding Morphological operators Discrimination (size and circularity) Euclidian distance transform
More 2011 [44]	SEM	Rat	Sciatic nerve	Morphological operators Thresholding Discrimination (size, circularity and intensity level)
Begin 2014 [33]	CARS	Mouse	Spinal cord	Extended-minima Active contours Discrimination (circularity and solidity) Edge detection Watershed
Zhao 2010 [48]	TEM	Rat	Optic nerve	Region growing Discrimination (size and circularity) Morphological operators
Weyn 2005 [42]	OM	Rat	Tibial nerve	Thresholding
Urso-Baiarda 2006 [43]	OM	Rabbit	Motor and facial nerves	Thresholding
Gierthmuehlen 2013 [47]	OM	Rat	Sciatic nerve	Region growing
Liu 2012 [45]	OM	Rat	Sciatic nerve	Thresholding (Otsu) Discrimination (size, circularity and solidity) Active contours
Fok 1996 [46]	TEM	Not specified	Not specified	Hough transform Active contours
Wang 2012 [60]	OM	Rat	Sciatic nerve	Watershed
Methods based on deep learning				
Naito 2017 [58]	OM	Human	Sural nerve	Clustering CNN (classification)
Mesbah 2016 [59]	OM	Mouse	Spinal cord	CNN (semantic segmentation)

After analyzing the work done on axon and myelin segmentation in the literature, several observations can be made:

- Several research groups use a combination of algorithms for the axon and myelin segmentation task.
- Most frameworks exploit the intensity/contrast difference between axons and myelin sheaths in microscopy images.
- Almost all research groups integrate an axon and/or myelin discrimination step based on size and circularity into the segmentation pipeline.
- Most frameworks are based on standard image processing algorithms. The reported methods that are based on neural networks are very recent (<2 years).
- Most pipelines are not fully automatic. They either require parameter tuning (e.g. set threshold values for axon candidate discrimination based on size) and/or postprocessing (e.g. manual correction). Therefore, segmentation results can vary across centers and research groups, limiting comparison between studies.
- All segmentation pipelines are designed to work on a specific microscopy modality. Consequently, they have only been validated on the modality for which they were designed for.
- Most research groups did not make the implementation available to the research community (only Begin and collaborators [33] shared the code online).

2.5 Axon and myelin morphometrics

Robust segmentation of axons and myelin sheaths from microscopy data would enable extraction of useful quantitative information about the microstructure of a tissue. **Figure 2.13** summarizes the most common axon and myelin morphometrics reported in the literature. A more in-depth review of the possible applications of morphometrics extraction is presented in section 2.6.

Individual metrics that can be computed from each axon-myelin pair include the axon diameter (inner diameter), the myelinated fiber diameter (outer diameter), the myelin thickness,

the axon area, the myelin area and the fiber g-ratio. In most cases, the equivalent axon diameter is reported, computed as the diameter of a circle with the same area as the axon. Axon diameter distributions of nervous tissues have been widely reported in previous studies. The caliber of a nerve fiber correlates with the signal transmission speed [61], and the severity of axonal damage often depends on the nerve fiber size in many neurological disorders [62].

The fiber g-ratio, which is defined as the ratio of the inner to the outer diameter of a myelinated nerve fiber, is also a relevant metric. It has been extensively used to quantify demyelination in tissues, in both microscopy [63]–[65] and MRI [64], [66], [67] studies. It has also been shown that the g-ratio can be estimated from the axon volume fraction (AVF) and myelin volume fraction (MVF) measurements [64]. This formulation of the *aggregate g-ratio* is particularly useful in quantitative MRI, as g-ratio measures can be estimated from MRI contrasts/techniques sensitive to axon and myelin content. In microscopic images, the aggregate g-ratio can be obtained by computing axon and myelin area for a given region. An interesting advantage of that formulation is that there is no need to explicitly compute the inner and outer diameters of each individual axon of the region of interest.

Finally, fiber density (axon density) is a metric that can be computed from the axon count of a region of interest in the sample. It can be used to assess nerve fiber compactness in different spinal cord white matter tracts [25].

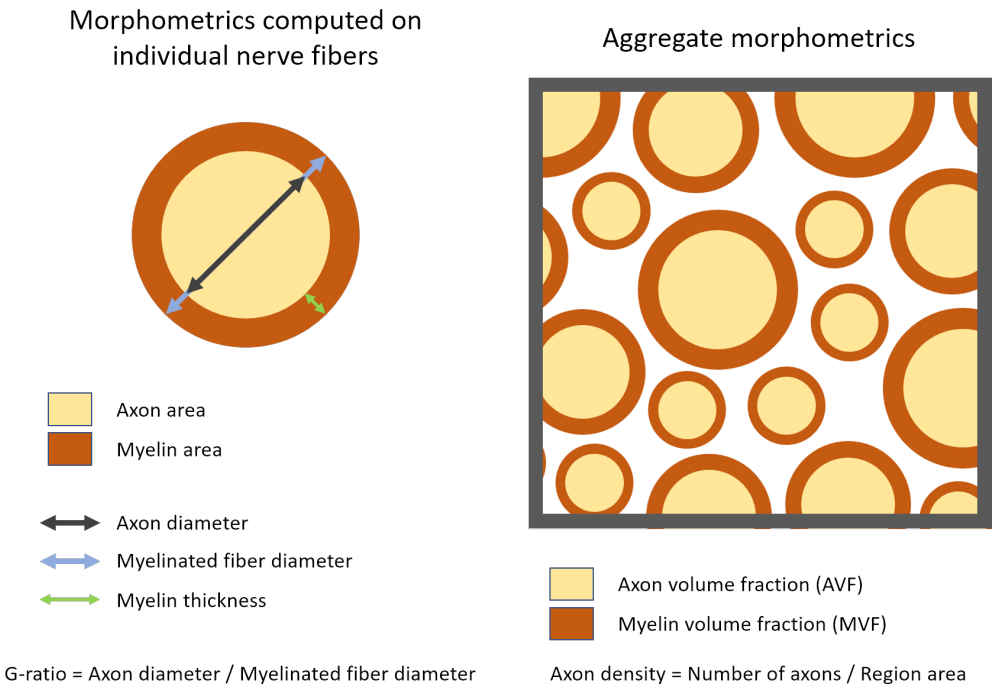

Figure 2.13: Schematic diagram illustrating the most common axon and myelin morphometric measures reported in the literature. The axon volume fraction (AVF) and myelin volume fraction (MVF) of a region of interest can be used to compute the aggregate g-ratio.

2.6 Applications of morphometrics extraction

The design and development of a robust axon and myelin segmentation pipeline capable of extracting morphometrics is strongly motivated by the ongoing research. Three major applications of nerve morphometry extraction are identified and discussed in the next sub-sections: (i) documentation of myelinated nerve morphometry across species and tissues; (ii) study of demyelinating models and (iii) validation of new MRI biomarkers.

2.6.1 Document nerve morphometry

There is an interest to document myelinated nerve morphometry across different species or tissues. Densities and counts of myelinated axons in the main spinal cord white matter tracts have been reported across different species (rat, cat, monkey and human) [25]. For instance, in the rat spinal cord, nerve fibers of the corticospinal tract are more densely packed and have smaller axon diameters [68]. Documenting the morphological differences between spinal cord tracts could lead to the construction of templates that could benefit the neuroscience community.

Documentation of nerve morphometry could also be useful when studying the effects of aging. For instance, Ugrenovic and collaborators [63] analysed the diameter and g-ratio of myelinated nerve fibers of the human sciatic nerve during the aging process. Their study identified a loss of large myelinated nerve fibers and a predomination of small fibers with thinner myelin sheaths in old subjects.

Other studies report average axon diameter and g-ratio values obtained from TEM samples of macaque corpus callosum [69] and mouse corpus callosum [65]. These studies could serve as reference for future investigations in same species and/or tissues.

2.6.2 Study demyelinating models

Several demyelinating animal models have been proposed to study diseases such as multiple sclerosis. Although none of these models can perfectly mimic the complexity of the human demyelinating diseases [70], they can be useful for studying some of the features of the disease.

Schultz and collaborators [71] study demyelination and remyelination in cuprizone-induced (mice) and focal lysolecithin-induced (rat) models. Li and collaborators [72] opt for a shiverer mouse model to study the effects of demyelination on corpus callosum strips. Similarly, Stidworthy and collaborators [73] aim at quantifying the early stages of remyelination in a cuprizone-induced model of mice. Mice corpus callosum samples are imaged with TEM and average g-ratio, myelin thickness, axon diameter and fiber diameter values are compared between normal and cuprizone-treated mice. West and collaborators [65] use average axon diameter and g-ratio values to compare normal and hypomyelinated mouse corpus callosum tissues.

Other studies aim at studying the relationship between traumatic brain injury and the demyelination of intact axons [74], evaluate the contribution of demyelination to the etiology of trigeminal neuropathic pain in rat samples [75], or conduct histological studies of sural nerves in patients with chronic inflammatory demyelinating polyneuropathy [76].

2.6.3 Validate new MRI techniques

MRI has been a central clinical tool for the imaging of the CNS. However, conventional MRI measures do not correlate well with patient disability. Having more sensitive and specific

measures would, for instance, help monitor the effect of remyelination treatments in MS. New quantitative MRI measures such as magnetization transfer ratio (MTR), diffusion and MR spectroscopic metabolite concentration [77] require extensive validation before being employed in a clinical context. One way of doing the validation would be to compare MRI metrics against measures obtained from high resolution microscopy (in this case, microscopy can be considered as a gold standard as it can capture individual axon-myelin morphometrics).

New methods have been developed to estimate the myelin g-ratio from MRI [78]. These myelin g-ratio measures can then be compared to the ones obtained from histology for validation [64]. Other methods aim at measuring the axon diameter distribution from diffusion MRI [79]. Kozlowski and collaborators [80] characterize white matter damage in the rat spinal cord using diffusion tensor imaging (DTI) and quantitative T2-weighted MRI and compare results to histology observations. Similarly, Mottershead and collaborators [81] explore different techniques (proton density, T1, T2, magnetization transfer ratio, diffusion weighted) to identify the ones that correlate the most with myelin content and axon density measures obtained from histology.

2.7 Research objectives

The literature review illustrated the wide range of methods in the application of axon and myelin segmentation of microscopy images. Limitations of the previous segmentation frameworks were also identified. Most frameworks are designed to work on a specific microscopy modality and have not been tested on other modalities. Moreover, the proposed pipelines are not fully automatic: either hand-engineered segmentation parameters or postprocessing tools are required. Also, the great majority of past frameworks are not open source and therefore not available to the research community.

The main goal of the axon and myelin segmentation is to extract morphological measures such as axon diameter and myelin thickness from the image samples. The relevance of the extraction of morphometrics from microscopy images was confirmed by identifying three major applications in the literature: documentation of nerve morphometry across species, tissues or tracts, study of the demyelination process in animal models, and validation of new quantitative MRI biomarkers.

The **general objective** of this research project is to *design, implement and validate an automatic axon and myelin segmentation framework for microscopy images and use it to extract relevant morphological metrics from the samples*.

The project can be divided into **3 specific objectives**:

1. Development of an automatic axon and myelin segmentation pipeline for microscopy samples of myelinated tissues.
2. Validation of the segmentation pipeline on commonly used microscopy modalities (optical microscopy, transmission electron microscopy, scanning electron microscopy, CARS microscopy), tissues (spinal cord white matter and brain corpus callosum) and species (rat, mouse, cat, human).
3. Implementation of a pipeline for morphometrics extraction, including axon diameter, myelin thickness, g-ratio and axon density.

An underlying goal of this project is to make the resulting software/implementations open source. The frameworks will be made publicly available via the *GitHub* platform. This means that special attention should be taken to ensure that the packages are user-friendly (i.e. provide relevant documentation on their usage, easy installation procedures, set up user support procedure). Moreover, the feedback and/or possible code contributions from users can strongly benefit this project in the long term.

CHAPTER 3 OVERALL METHODOLOGY

This chapter presents an overview of the methodology used in the project. The methodology was guided by the three specific objectives identified in the previous chapter. First, two segmentation frameworks were designed and implemented: *AxonSeg* and *AxonDeepSeg*. *AxonSeg* is based on a combination of standard image processing algorithms, while *AxonDeepSeg* is based on the recent advances in neural networks and proposes a convolutional neural network architecture. Validation of both frameworks was performed on common microscopy modalities, tissues and species, against manually labeled data. A pipeline to extract morphological measurements from the segmentation outputs was developed for both frameworks and applied to extract mean morphometrics from spinal cord white matter tracts. The method for each package is detailed in the next chapters.

3.1 Resulting publications

The work done during this research project has led to the publication of two articles in peer-reviewed journals. The first paper presents the *AxonSeg* segmentation framework and has been published in *Frontiers in Neuroinformatics* (Impact Factor = 3.9). The second paper introduces the *AxonDeepSeg* segmentation framework and has been published in *Nature Scientific Reports* (Impact Factor = 4.3). Both papers were published in open source journals, which is in accordance with the nature of the software packages developed and the philosophy of the *NeuroPoly* lab (both *AxonSeg* and *AxonDeepSeg* are open source and available on *GitHub*). This project was also presented during a poster session at the *10th QBIN (Quebec Bio-imaging network) Scientific Day* (poster #16).

The methods developed in the context of this project also contributed to other projects in the lab. For instance, the segmentation frameworks developed were used to automatically segment full slice spinal cord samples of 5 rats across all 34 segments (i.e. approximately 175 spinal cord cross sections). Morphometrics maps obtained from the large-scale segmentation were used to generate a template of the white matter microstructure of the rat spinal cord. A paper is currently being written.

Chapter 4 presents *AxonSeg*, which is the topic of the first paper written during this project. The steps of the axon and myelin segmentation pipeline are detailed. The preliminary

axon segmentation is done by the extended-minima algorithm, and a discriminant analysis algorithm based on both shape and intensity features is used to eliminate false positives. The preliminary myelin segmentation is done by an edge detection of the boundaries. Then, axon and myelin segmentations are updated by using a double active contour algorithm (for both boundaries of the myelin sheath). The pipeline was validated against manual segmentations on CARS, OM and SEM microscopy using the Dice similarity metric and nerve fiber sensitivity and precision metrics. The morphometrics extraction pipeline is demonstrated by applying *AxonSeg* to a full cat spinal cord slice and extracting mean axon and myelin measures from regions of interest.

Chapter 5 presents the second paper, which focuses on the *AxonDeepSeg* software package. The convolutional neural network architectures designed and optimized for the SEM and TEM models are explained, along with the training procedure (data augmentation, dropout, batch normalization, separation in patches). Validation of both SEM and TEM models is done on unseen test data against manual segmentations, using the axon and myelin Dice similarity metrics, the 3-class pixel-wise accuracy and the sensitivity and precision measures. Note that the 3-class pixel-wise accuracy is added here as a validation metric. This accuracy is also used in the training procedure as metric of the cost function and is commonly used in deep learning validation pipelines involving multi-class pixel classification. The morphometrics extraction pipeline used is also detailed and validated by extracting tract-based measures and comparing the obtained metrics with known morphology.

Chapter 6 presents supplementary methods and results, notably the robustness of *AxonDeepSeg* models to noise/blurring/intensity changes, the effect of the patch overlap value on the *AxonDeepSeg* prediction, new performance metrics for the validation of the *AxonDeepSeg* segmentation, and a comparison between *AxonSeg* and *AxonDeepSeg* performances.

CHAPTER 4 ARTICLE 1: AXONSEG: OPEN SOURCE SOFTWARE FOR AXON AND MYELIN SEGMENTATION AND MORPHOMETRIC ANALYSIS

This paper presents the *AxonSeg* framework designed for axon and myelin segmentation of microscopy images. The segmentation pipeline is based on a combination of intensity, edge and energy minimization algorithms. A discriminant analysis model combining both shape and intensity features is proposed for the axon discrimination step. The software package offers a graphical user interface (GUI) to help set the segmentation parameters for a given microscopy modality/tissue. The segmentation pipeline is validated on SEM, OM and CARS samples from rat and cat spinal cords. Application to large-scale segmentation and morphometrics extraction is demonstrated.

My **contributions** include the design and implementation of the axon segmentation, axon discrimination, morphometrics extraction, GUI and segmentation validation modules of *AxonSeg*, as well as the generation of the results, writing of the paper and the generation of the figures. This paper was published in *Frontiers in Neuroinformatics*.

doi: 10.3389/fninf.2016.00037.

Authors: Aldo Zaimi[1], Tanguy Duval[1], Alicja Gasecka[2,3], Daniel Côté[2,3], Nikola Stikov[1,4] and Julien Cohen-Adad[1,5].

Affiliations:

[1]Institute of Biomedical Engineering, Polytechnique Montreal, Montreal, QC, Canada;

[2]Institut Universitaire en Santé Mentale de Québec, Quebec, QC, Canada;

[3]Centre d'Optique, Photonique et Laser, Université Laval, Quebec, QC, Canada;

[4]Montreal Heart Institute, Montreal, QC, Canada;

[5]Functional Neuroimaging Unit, CRIUGM, Université de Montréal, Montreal, QC, Canada.

4.1 Abstract

Segmenting axon and myelin from microscopic images is relevant for studying the peripheral and central nervous system and for validating new MRI techniques that aim at quantifying tissue microstructure. While several software packages have been proposed, their interface is sometimes limited, and/or they are designed to work with a specific modality (e.g., scanning electron microscopy (SEM) only). Here we introduce *AxonSeg*, which allows to perform automatic axon and myelin segmentation on histology images, and to extract relevant morphometric information, such as axon diameter distribution, axon density and the myelin g-ratio. *AxonSeg* includes a simple and intuitive MATLAB based graphical user interface (GUI) and can easily be adapted to a variety of imaging modalities. The main steps of *AxonSeg* consist of: (i) image pre-processing; (ii) pre-segmentation of axons over a cropped image and discriminant analysis (DA) to select the best parameters based on axon shape and intensity information; (iii) automatic axon and myelin segmentation over the full image; and (iv) atlas-based statistics to extract morphometric information. Segmentation results from standard optical microscopy (OM), SEM and coherent anti-Stokes Raman scattering (CARS) microscopy are presented, along with validation against manual segmentations. Being fully-automatic after a quick manual intervention on a cropped image, we believe *AxonSeg* will be useful to researchers interested in large throughput histology. *AxonSeg* is open source and freely available at: https://github.com/neuropoly/axonseg.

Keywords: axon, myelin, segmentation, discriminant analysis, histology, microscopy, graphical user interface, g-ratio.

4.2 Introduction

The central nervous system, which consists of the brain and the spinal cord, relies on the transmission of electrical signals via white matter axons. The myelin sheath, wrapped around the axons, has a key role in the transmission process over long distances (Zoupi et al., 2011; Seidl, 2014). In case of degenerative diseases such as multiple sclerosis, myelin tends to degenerate by undergoing significant morphological changes, affecting signal propagation (Lassmann, 2014; Alizadeh et al., 2015; Papastefanaki and Matsas, 2015). A large body of research focuses on the

understanding of the intrinsic patterns related to demyelination in animal models (e.g., Experimental Autoimmune Encephalomyelitis, shivered, Wallerian degeneration) (Baker and Amor, 2014; Ben-Nun et al., 2014; Papastefanaki and Matsas, 2015). Therefore, it is of particular interest to image white matter microstructure with high enough resolution to identify axon and myelin morphology. Histology has provided valuable information, but popular imaging techniques such as transmission electron microscopy (TEM) can only image small regions (typically $\sim 100 \times 100 \mu m^2$). New imaging techniques with a sliding acquisition window and stitching capabilities have emerged that can provide a full picture of a sample under investigation, e.g., a $1 \times 1 cm^2$ cross-section of a spinal cord. However, axon and myelin segmentation of these large datasets is extremely time-consuming and difficult, as a dataset covering several cm^2 contains millions of axons. Moreover, manual segmentation is subject to user bias and is therefore not reproducible within and across sites.

The first software tools capable of accomplishing segmentation of nerve fibers have mostly focused on simple segmentation algorithms for microscopic images stained with toluidine blue (Cuisenaire et al., 1999; Romero et al., 2000). Some research groups have opted for manual segmentation (Berthold Ch Fau - Nilsson et al., 1983; Dula et al., 2010; Liewald et al., 2014): Lieward and collaborators performed manual segmentation on TEM samples in order to study the distribution of axons diameters in the cortical white matter. Begin et al. (Begin et al., 2014) introduced an algorithm capable of segmenting both axon and myelin in large-scale images from coherent anti-Stokes Raman scattering (CARS) microscopy. Their software can be used to extract morphological data from the input images and is fully automated. Although their software could be used on different contrasts, all the parameters are set to work on CARS images, and no GUI is included to adapt them to other contrasts. More et al. (More et al., 2011) introduced a simple semi-automated algorithm designed to work on scanning electron microscopy (SEM) images. While a GUI is included, its function is limited to segmentation parameters only, which are hard to provide without prior knowledge of the tissue sample to segment. Moreover, the axon segmentation needs to be manually corrected before launching the myelin sheath segmentation and there is no integrated framework for extracting relevant morphometric information afterwards. Other segmentation tools focus on TEM images from optic nerves (Zhao et al., 2010) or cross-sectional images of rat nerve fibers from sciatic nerve in optical microscopy (OM) (Wang et al., 2011). However, both work only on specific imaging modalities.

Most of the published work describes the segmentation algorithms without giving open access to the related scripts, or without providing an intuitive interface for other researchers to use. In summary, there is no single software for axon/myelin segmentation that is adapted to all imaging contrasts, is available for free, handles large-scale histology data and has a graphical user interface (GUI). Having such software would facilitate the processing of large microscopy images and standardize processing across research groups.

In this paper, we introduce *AxonSeg*, which is designed to perform axon and myelin segmentation on large-scale histology images, features an intuitive GUI, works with various contrasts and is open source. This paper is organized as follows: (i) the Methods section lists the main steps of *AxonSeg*, details the segmentation strategies and the discrimination model, then details the validation part, (ii) the Results section presents validation results, proof-of-concept axon and myelin segmentation obtained from three different contrasts (CARS, OM and SEM) and shows statistics results on relevant morphological metrics from the input images, (iii) the Discussion section addresses pros/cons of *AxonSeg* and discusses further possible developments, and (iv) an example use case describes the typical workflow in order to segment an optical microscopy sample.

4.3 Materials and Methods

4.3.1 Algorithm

AxonSeg aims at performing both axon and myelin segmentation on various imaging contrasts, including a robust axon candidate discrimination step that aims at optimizing sensitivity and precision.

4.3.1.1 Axon segmentation

The axon segmentation strategy is based on Begin et al. (Begin et al., 2014) and uses the extended minima method (also known as gradient-based region growing method) to output binary segmentations of the intracellular part of the axon (i.e., axon without myelin). The extended minima algorithm is defined as the regional minima of the H-minima transform. The H-minima transform eliminates all minima whose intensity is less than input threshold h. The regional

minima, i.e. connected components of pixels with the same intensity value, and whose outer boundary pixels have higher values, are then computed.

The binary axon segmentation image is then post-processed by common morphological operations: remove isolated individual pixels, fill isolated interior pixels, perform morphological closing, remove H-connected pixels, perform morphological opening and finally, remove all partial axon candidates that touch the outside border.

4.3.1.2 Axon discrimination

During the axon segmentation step, false positives are inevitably introduced in the resulting output image. Thus, a discrimination step is needed to keep most of the true axons while trying to reject the false ones. Our discrimination strategy aims at combining shape (morphology) and intensity features in order to build a discriminant analysis (DA) classifier that distinguishes true/false axons.

The DA model is initiated by using a training dataset as input. This training dataset is generated by the user using the procedure described in section “Discriminant analysis” and is made of two groups: one group contains true axons and the other group contains false axons. Shape features are determined for every labeled object of the training groups:

- Circularity: describes the roundness of the object, defined as $4\pi \times \text{Area}/\text{Perimeter}^2$.
- Solidity: describes the compactness of the object, computed as Area/ConvexArea, where the convex area is the area of the polygon containing the object.
- Ellipticity: describes the flattening of the object, defined as the ratio between the minor and major axes.
- Equivalent diameter: diameter of the object, computed as the diameter of a circle with the same area as the object.

In a similar way, intensity features are computed for every labeled object of the training groups:

- Object intensity mean and standard deviation.
- Neighbourhood intensity mean and standard deviation: the neighbourhood is defined by performing a small morphological dilation of the object (disk-shaped structuring element with a radius of 2 pixels).

- Contrast: intensity difference between object and neighbourhood intensity means.

True axons are usually described as round, convex, low intensity shapes, enclosed by a higher intensity myelin sheath annulus. The output of this process is a linear or quadratic classifier in the n-parameters space that can predict true/false axons from the untrained dataset.

4.3.1.3 Myelin segmentation

The myelin segmentation strategy is based on the algorithm developed by Begin et al. (Begin et al., 2014) which relies on radial screening of the axon neighbourhood and the minimal-path algorithm (Vincent, 1998). First, after labeling the axons, the gradients of the radial profiles of each axon are computed by using a Sobel filter. The minimal-path algorithm is applied on radial profile gradients in order to detect the outer border of the myelin. In our implementation, we apply a subsequent maximal path to detect the inner border of the myelin sheath in order to refine the axon segmentation. Also, additional constraints have been added in order to improve the robustness of the myelin segmentation. This was done using a double snake algorithm (active contours) adapted from MATLAB Central File Exchange[1], designed to detect both the outer and the inner boundary of the myelin sheath, and constrained to have a homogenous myelin thickness across the axon circumference (More et al., 2011) and a g-ratio (defined as the ratio of the inner to the outer diameter of the myelin sheath) within the range 0.4 and 1.

Next, a cleaning step is done by verifying the presence of conflicts between adjacent myelin areas. If more than 50% of the myelin sheath area from one axon is overlapping with the myelin sheath area from another adjacent axon, then the former axon is rejected from the analysis. The final axon segmentations are obtained from the corresponding myelin segmentations after computing a morphological filling.

4.3.2 AxonSeg software

Figure 4.1 shows the main steps in *AxonSeg*: cropping of a small region, image pre-processing, axon segmentation and discriminant analysis, and axon and myelin segmentation of

---

[1] Kroon, Dirk-Jan (2010). Snake: Active Contour (http://www.mathworks.com/matlabcentral/fileexchange/28149-snake---active-contour), MATLAB Central File Exchange. Retrieved October 30, 2015.

the full image. The main GUI tool of *AxonSeg* can be accessed by launching the *SegmentationGUI* function. Note that if the image is already small (e.g., less than 2000×2000 pixels), it is possible to bypass the cropping step and the discriminant analysis step and just run the segmentation over the full image. Each step has a previewing capability and is composed of a "Go to next step" button that uses the selected parameters/options to generate and display the output. A "Reset Step" button allows the user to go back to the previous step.

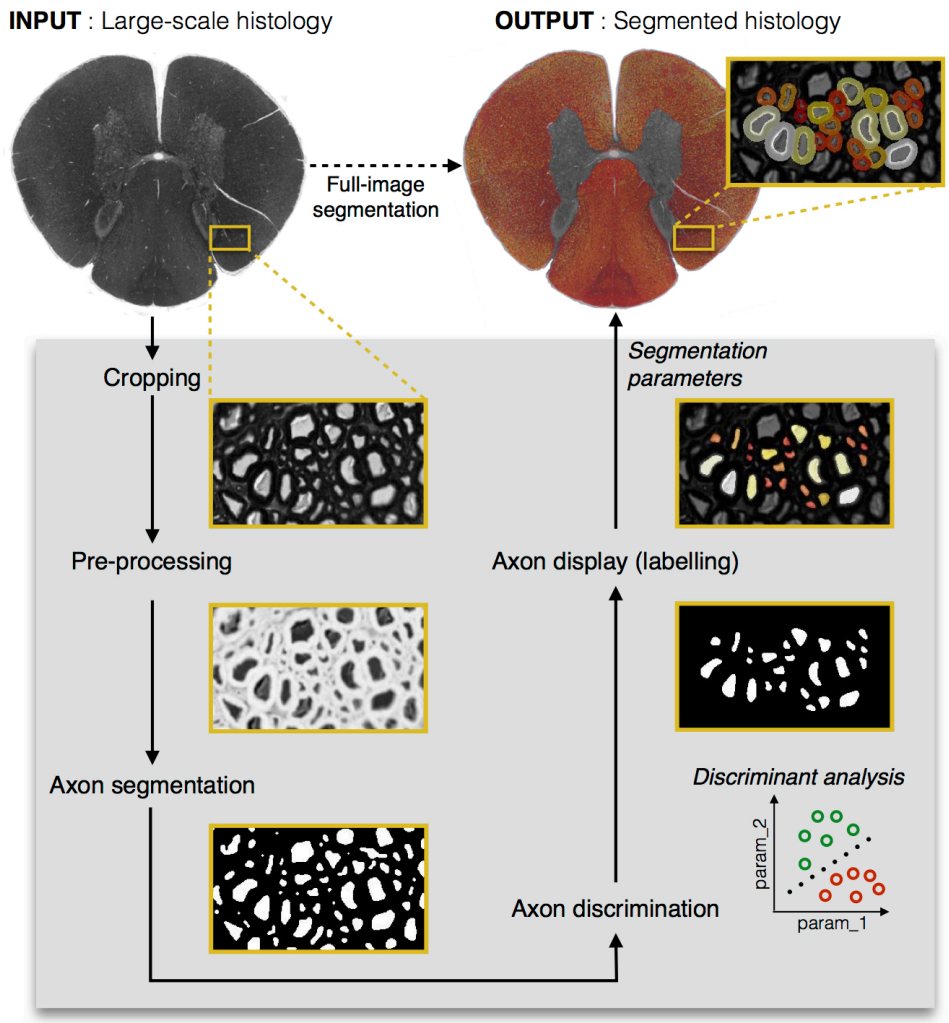

Figure 4.1: Diagram illustrating the main steps of *AxonSeg*. Pre-processing, axon segmentation and discrimination are first performed on a cropped image. Then, the segmentation parameters applied on the cropped image are saved and can be used to launch the automatic full-scale segmentation of the axons and the corresponding myelin sheaths. Note that axons within a range of 2 µm from the edge of the image are discarded to prevent the segmentation of incomplete axons.

4.3.2.1 Pre-processing of cropped image

At the beginning of the process, the user has access to the following pre-processing tools: (i) a colour inversion (complement) module to ensure that axons are darker than the myelin, as required by the segmentation strategy; (ii) a smoothing module (averaging filter, size 3×3) and (iii) contrast enhancement via histogram equalization and deconvolution. The user is asked to enter the pixel size, so that all morphological measurements regarding axon or myelin are given in distance units. If a scaling bar is present on the image (sometimes integrated from microscope outputs), the user can define the size of the scaling bar and select the two extreme points of this bar to get the pixel size.

4.3.2.2 Axon segmentation from cropped image

Axon segmentation is performed using the extended minima algorithm, as explained in section "Axon Segmentation". The user can adjust segmentation results in real time by tuning the parameters of the extended minima (threshold values) with a slider (see "Example Use Case" section for a typical use case).

4.3.2.3 Discriminant analysis

A first global discrimination step, conducted by the user, aims at cleaning up the axon candidates field by eliminating a significant number of false positives. The "Minimal size" slider is used to remove false positives associated with small debris or artifacts (e.g., holes naturally present in the tissue), while the "Solidity" and "Ellipticity" sliders can help eliminate false positives coming from inter-nerve-fiber regions (extended intercellular spaces) of the image. The "Go to next step" button combines the result of each slider. On the next step, the user can manually remove the remaining false positives by clicking on the image. The selection of the features for the prior axon discrimination was made in accordance with the parameters analysis performed on control datasets. Similar feature analysis was performed by other research groups in order to find the best parameters that can separate the true and false positives (Romero et al., 2000; Zhao et al., 2010; More et al., 2011; Wang et al., 2011; Begin et al., 2014), but none of them were able to find a perfect combination due to the large variability among axon shapes and intensities.

The discriminant analysis tool is then launched on the corrected cropped image: the remaining axons are considered as true positives while those eliminated earlier are considered as false positives. The user can select either the linear discriminant analysis or the quadratic one: the classification results will be displayed on the GUI, and the user can scroll through the "Discriminant analysis" slider in order to select the sensitivity and specificity combination adapted to their needs. Opting for a higher sensitivity is a way to keep more axons, although it usually leads to lower specificity, thus accepting more false positives in the axon segmentation output. In our application, we aim at obtaining a classifier with maximal sensitivity and maximal specificity, i.e. the closest possible to the upper left corner of the Receiver Operating Characteristic (ROC) curve. The minimal Euclidean distance from the upper left corner can thus be computed and is available in the GUI along with other options (e.g. maximal sensitivity, maximal specificity). **Figure 4.2** shows a screen capture of the GUI during the DA step.

Note that we do not let the user add missing axons in the GUI at this step for two main reasons. First, manually adding missed axons can add bias in shapes and thus strongly affect the discriminant analysis. Secondly, *AxonSeg* comes with a smaller GUI tool called *ManualCorrectionGUI*, which can be used in order to correct a segmentation result by adding, removing or modifying axons after the full-image segmentation.

4.3.2.4 Myelin segmentation on cropped image

The myelin segmentation is performed on the accepted axons (after manual correction or as determined by the selected discriminant analysis result), as explained in section "Myelin Segmentation". The results are displayed on the GUI. The user can go back to previous steps and can adjust the parameters iteratively until satisfied by the final result.

4.3.2.5 Outputs

At the end of the discriminant analysis and myelin segmentation on the cropped image, an output folder is created. A very important feature of our software is the *axonlist* structure, which stores for each axon object the following fields: the data (all the pixels belonging to the axon object, in x and y coordinates), the axon and myelin areas (in both pixels and μm^2), the centroid (in x and y coordinates), the axon ID (from labeling), the myelin g-ratio, the axon equivalent diameter (in μm), the myelinated fiber equivalent diameter (in μm) and the myelin thickness (in

µm). These metrics are also stored in a comma-separated values (CSV) file, also in the output folder. **Figure 4.3** illustrates the main morphological metrics computed by *AxonSeg*.

The default segmentation images saved in the output folder are labeled in axon equivalent diameter color code: a first one displaying the axon segmentation and a second one displaying the myelin segmentation. Other segmentation displays can be produced by using the *as_display_label* function included in *AxonSeg*. Two display types are available (either axon or myelin), color mapped in axon equivalent diameter, g-ratio, myelin thickness or axon ID. Additional images are also saved in the output folder: pre-processed input cropped image, binary image of the initial axon segmentation candidates and binary image of the final axon segmentation result after discrimination.

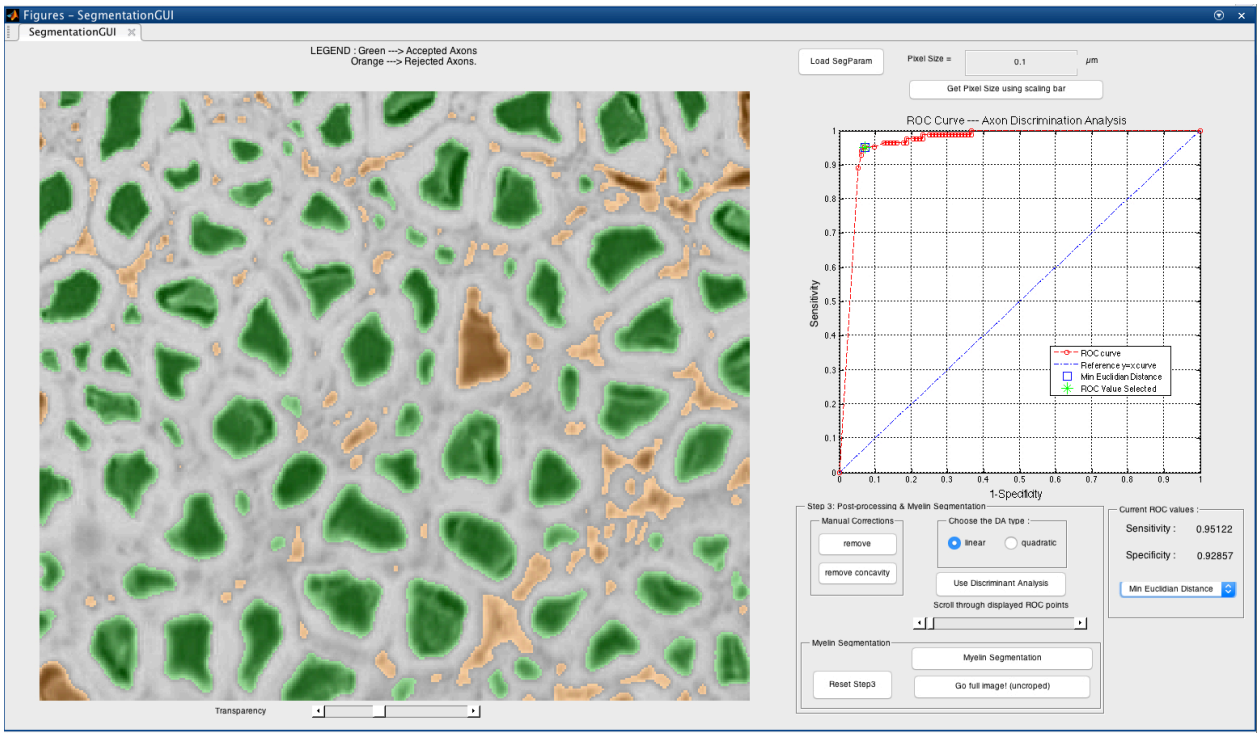

Figure 4.2: Discriminant analysis step in the *AxonSeg* GUI. The user can scroll through available sensitivity and specificity combinations and display the results. For each sensitivity/specificity value, accepted axons are displayed in green while rejected axons are displayed in orange. The user can also decide to select a ROC metric (e.g. minimal Euclidian distance, maximal sensitivity, maximal specificity). When satisfied with the discriminant analysis classifier, the user can launch the myelin segmentation.

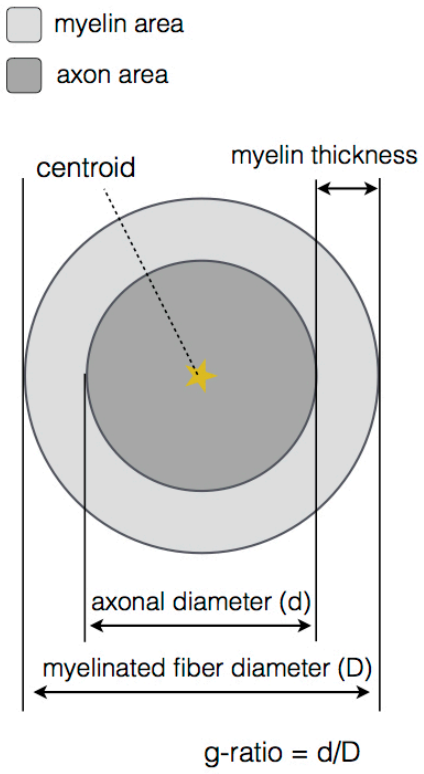

Figure 4.3: Main morphological properties computed by *AxonSeg*. For each myelinated fiber, the centroid, the axon and myelin areas, the axonal diameter, the myelinated fiber diameter, the myelin thickness and the myelin g-ratio are calculated.

All parameters that have been adjusted using the GUI are saved into the MATLAB structure *SegParameters.mat* which is subsequently used for the full image segmentation. This structure can also be loaded in the GUI at any time if needed, in particular if the user wants to segment another image with similar contrast. Relevant parameters include pre-processing options, axon segmentation thresholds, prior axon discrimination values and DA classifiers.

4.3.2.6 Full image segmentation

If the user worked on a cropped region of the initial image, the full image segmentation can be launched by applying the same parameters as those used throughout the segmentation and processing of the cropped region. To avoid RAM saturation, the full image is processed by smaller blocks, with 20% overlap, and then stitched together. Note that this feature allows the segmentation of much larger images than initial implementation by Begin et al. (Begin et al., 2014) (images up to 21000×12000 pixels were processed successfully). This segmentation per

block also allows parallelization across CPUs by using the Parallel Computing Toolbox in MATLAB (*parfor* and *parpool*), if supported by the computer.

This full segmentation can be run without using the GUI, by calling the *as_Segmentation_full_image* function, which requires the image input name and the *SegParameters.mat* file (see "Example Use Case" section for an example). The block size, the overlap between adjacent blocks and the output folder name can also be specified if needed.

4.3.2.7 Morphometric analysis

In order to analyse the results, additional tools have been developed to extract statistics in specific regions of interest (ROI), for instance the axon diameter distributions in the posterior fasciculus of the spinal cord. Other statistics are: axon and myelin areas, myelin thickness, g-ratio and axon count. These ROI can be manually drawn or can be imported from a digital version of an existing atlas (e.g., the human white matter atlas from (Levy et al., 2015)). A registration module which is based on an affine 2D transformation is also provided to register the ROI to the segmentation. These operations (mask registration, labeling and metrics calculations) can be performed by a set of *AxonSeg* functions (see "Example Use Case" section for more details).

4.3.3 Validation

4.3.3.1 Data

Spinal cord images were acquired with three different imaging techniques: OM, SEM and CARS. Standard OM images were obtained from one rat and one cat (cervical sections). Samples were embedded in paraffin and imaged using a whole slide scanner with 20× magnification (Hamamatsu NanoZoomer 2.0-HT). SEM images were obtained from one rat (cervical section). Sample was stained in osmium, embedded in epoxy, polished and imaged using an SEM system (Jeol 7600F) with 1000× magnification (pixel size of 0.08 µm). CARS images were obtained from one rat (thoracic section). Sample was imaged with a 60× objective lens (UPLSAPO 1.2 NA w, Olympus) and recorded images were stitched to reconstruct the whole section (~0.2 µm/pixel). All animals were perfusion-fixed using 2% PFA and 2% Glutaraldehyde.

4.3.3.2 Ground truth

Ground truth images of axon segmentations were produced by manually correcting results of *AxonSeg*, using the *ManualCorrectionGUI* tool. Correction included: adding missed axons, removing false positives and correcting axons shape when necessary. The resulting binary images (identifying axons as logical true and background as logical false) were used to assess segmentation quality, sensitivity and precision.

4.3.3.3 Sensitivity and precision of axon detection

In order to evaluate the ability of *AxonSeg* to distinguish between true and false axon candidates, sensitivity and precision measurements were computed. True positive, false negative and false positive counts were obtained by automatically comparing the binary test and control images by using the centroid positions. Here, a true positive (TP) is defined as a correctly identified axon (present in both test and control images), while a false negative (FN) is defined as an incorrectly rejected one (present in control image but absent from the test image). We can also identify the false positives (FP), described as axon candidates present in the test image, but absent from the control image. The sensitivity, also called true positive rate (TPR), is defined as:

$$TPR = \frac{TP}{TP + FN} \tag{4.1}$$

The precision, also called positive predictive value (PPV), can be defined as:

$$PPV = \frac{TP}{TP + FP} \tag{4.2}$$

4.3.3.4 Quality of axon segmentation

The quality of the axon segmentations was measured by comparing segmentation results obtained from the GUI to the ground truths by using the Dice coefficient. Given two binary images I and J of the same size, we can define a as the number of pixels where the corresponding values of I and J are both 1 (true). In a similar way, we can also define b and c as the number of pixels where a 0 (false) value in I has a corresponding 1 value in J, and where a 1 value in I has a corresponding 0 value in J, respectively. We can then define the Dice coefficient between I and J:

$$D = \frac{2a}{2a + b + c} \tag{4.3}$$

For every axon object in test image *I* and its corresponding one in the ground truth image *J* (i.e. for every true positive detected), the Dice coefficient was calculated. 10^{th}, 50^{th} and 90^{th} percentiles were obtained from the Dice distributions. Note that we have not quantified the quality of the myelin segmentation, as this has already been done by Begin et al. (Begin et al., 2014).

4.4 Results

4.4.1 Sensitivity and precision of axon detection

To validate the sensitivity of axon detection (true positives), three conditions were tested: axon segmentation without DA, axon segmentation with linear DA and axon segmentation with quadratic DA. The axon segmentation without DA was performed by visual assessment on a cropped region, using the feature sliders available in the GUI and trying to keep most of the true positives while eliminating as much false positives as possible. Then, linear and quadratic DAs were computed by using the same cropped region as training dataset and selecting the sensitivity/specificity value for which the Euclidean distance metric was minimal. Cropped regions of about 25% of the full image were used in all cases. Sensitivity and precision were then computed on the full images. Results are reported in **Table 4.1** for the three available modalities (OM, SEM and CARS).

4.4.2 Quality of axon segmentation

Quality of axon segmentation was assessed by computing the Dice coefficient between the automatic segmentation and the ground truth (segmentation with manual correction). For all modalities tested, segmentation results were produced without the use of DA, as only true positives are considered in the Dice computation and because the use of DA does not affect the quality of the segmentation (it only affects the detection of axons). Results are reported in **Table 4.1** for the three modalities (OM, SEM and CARS).

Table 4.1: Assessment of axon detection and segmentation quality provided by *AxonSeg*. Three modalities were tested: OM, SEM and CARS. Sensitivity and precision were computed from the true positive, false positive and false negative counts for three conditions: no discriminant analysis (DA) model, linear DA model or quadratic DA model. Segmentation quality was assessed by using the true positives without DA (see "Quality of Axon Segmentation" section for justification): 10^{th}, 50^{th} and 90^{th} percentiles were computed from the Dice coefficient distribution.

		Axon Detection		**Segmentation quality (from individual axon Dice coefficients)**		
		Sensitivity	**Precision**	**10^{th} Percentile**	**50^{th} Percentile**	**90^{th} Percentile**
OM	**No DA**	0.8900	0.8194	0.7781	0.8636	0.9254
	Linear DA	0.8720	0.8519			
	Quadratic DA	0.8857	0.8304			
SEM	**No DA**	0.7886	0.6745	0.6876	0.8271	0.9221
	Linear DA	0.8607	0.5492			
	Quadratic DA	0.8593	0.5959			
CARS	**No DA**	0.5634	0.4751	0.7708	0.9234	0.9688
	Linear DA	0.5746	0.5126			
	Quadratic DA	0.5618	0.5181			

4.4.3 Application to large-scale image

Figure 4.4 shows results of full image segmentation (axon display) from a cervical cross-section of a cat spinal cord from OM, including a zoom in on a small region (myelin display). The displayed spinal cord slice has an area of about 4 mm^2. Visual assessment suggests fairly good segmentation quality, with the majority of axons detected and correctly segmented. This result can also be found on our laboratory website with a zooming feature[2]. Some axons are missed, which is largely due to the poor resolution of the image.

[2] http://www.neuro.polymtl.ca/histology.php

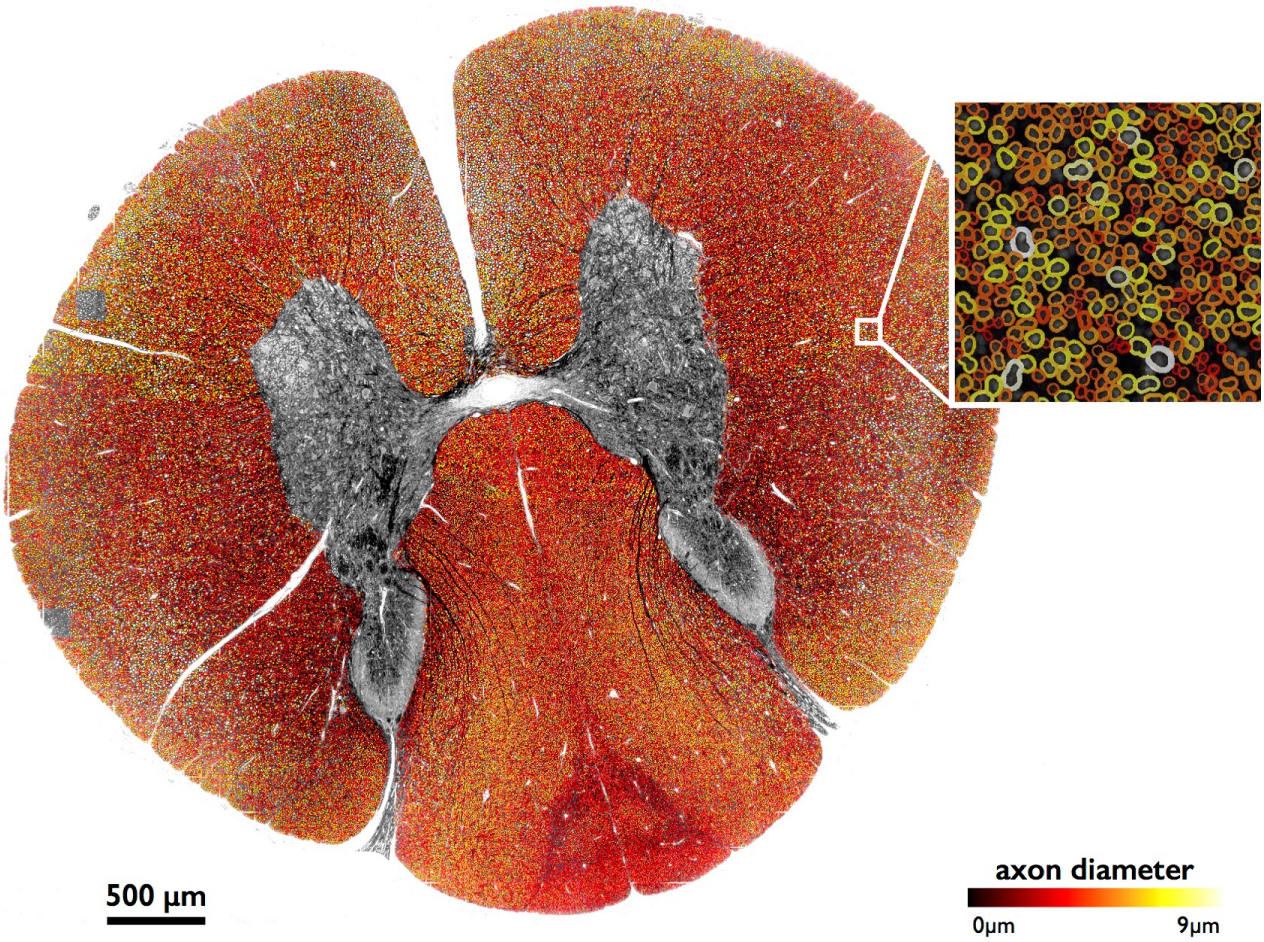

Figure 4.4: Large-scale segmentation (axon display) in a cat spinal cord (cervical section), color-coded for axon diameter. Myelin display (color-coded for axon diameter) is used in the zoomed region. Note that the regions outside of the white matter were masked out for better clarity, using automatic tools from *AxonSeg*.

4.4.4 Atlas-based morphometric analysis

Morphometric statistics were extracted from the same full-scale image as that in **Figure 4.4**. Results are shown in **Figure 4.5**. For each region, the following metrics were extracted: axon count, axon diameter and myelin g-ratio.

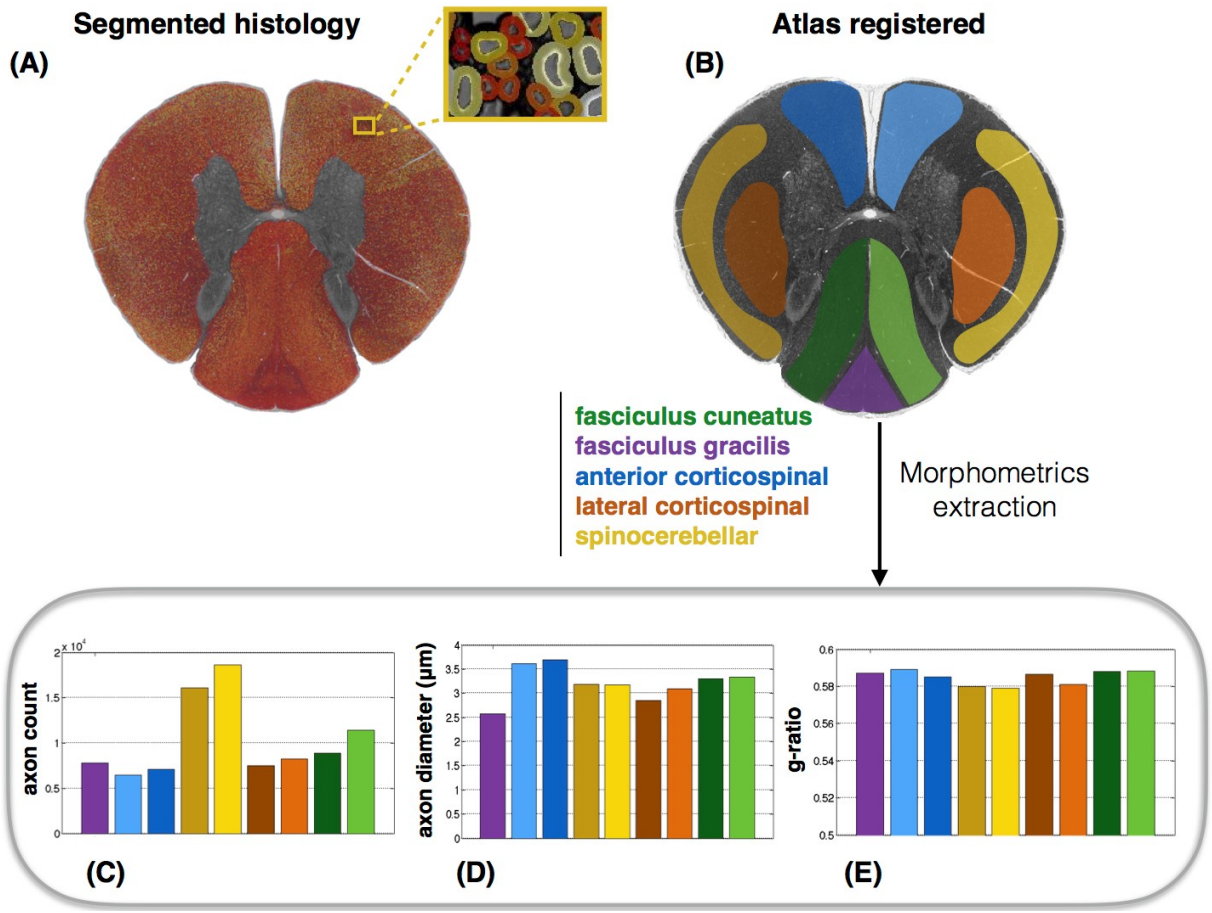

Figure 4.5: Atlas-based morphometry from segmented images. (A) Segmented histology results. (B) Atlas of white matter tracts registered to the histology image. Morphometric statistics for each of the tracts are extracted: (C) axon count; (D) axon diameter (mean); (E) myelin g-ratio. Note the slightly lower g-ratio compared to the expected values from the literature (about 0.7) (Chomiak and Hu, 2009), which is likely due to the poor resolution of the optical microscope inducing an over-segmentation of the myelin sheath.

4.4.5 Results in various imaging modalities

As a proof of concept, the software has been tested on three different histology contrasts: OM, SEM and CARS microscopy. **Figure 4.6** shows results of axon and myelin segmentation, displayed using an axon diameter colormap.

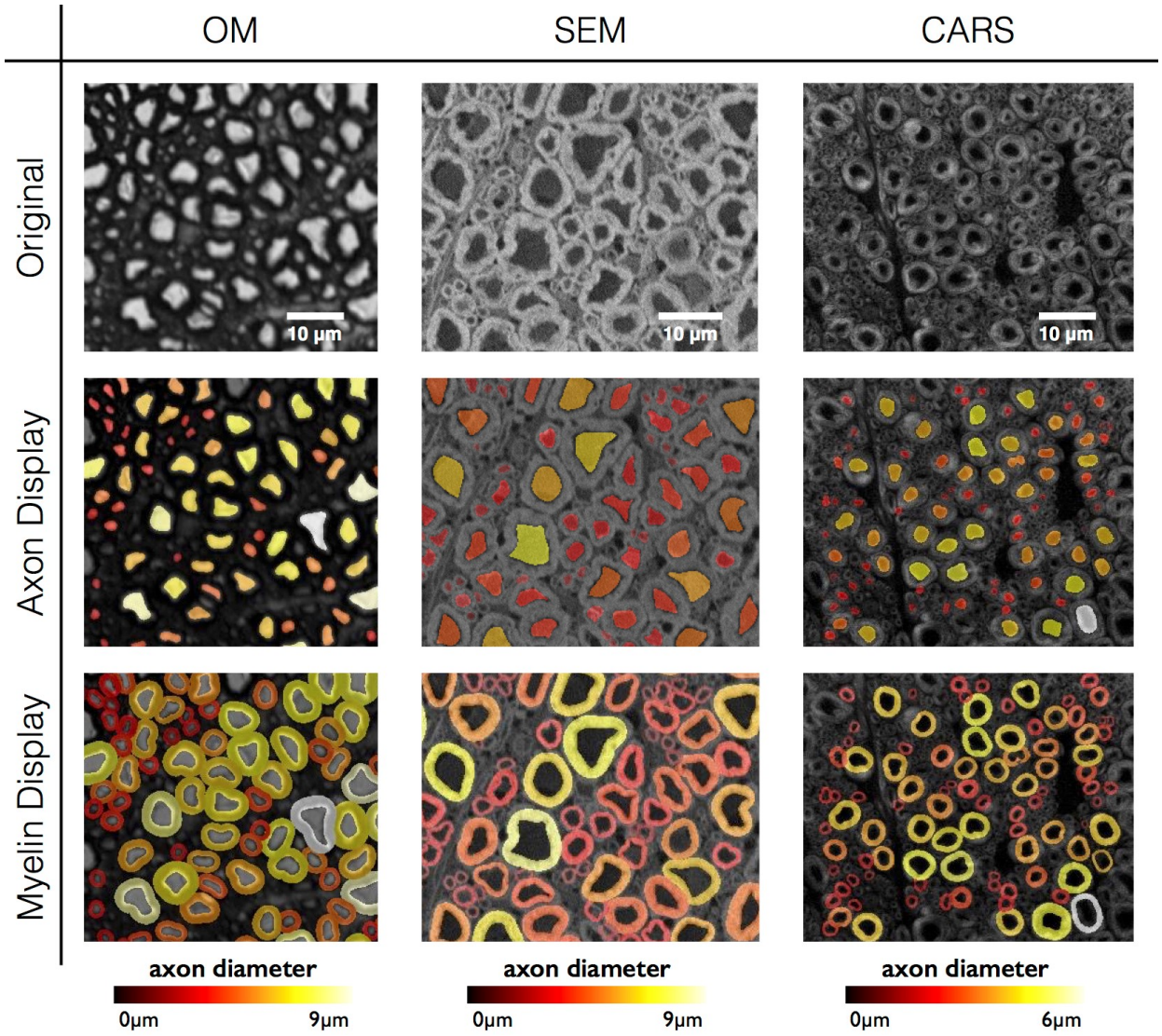

Figure 4.6: Results in various imaging modalities: optical microscopy (OM) from a cat spinal cord, scanning electron microscopy (SEM) and CARS microscopy from a rat spinal cord. All figures are color-coded for axon diameter. Note: myelin sheaths overlap more on the OM contrast due to the low resolution, inducing a blurring of the myelin sheath and therefore an apparent over-segmentation. This effect could be compensated using deconvolution algorithms applied to the image, although this will require further investigation.

4.4.6 Computation time

All data presented in this article were processed on a Macintosh (2.9 GHz Intel i5 processor, 4 cores, 8 GB 1600 MHz DDR3 RAM). The average computing time to process a single axon (including both axon and myelin segmentation) was 0.26 s. Therefore, it takes less

than 8 hours to automatically process a large-scale dataset (e.g. 21000×12000 pixels image, around 100,000 axons).

4.5 Discussion

This paper introduced the Axon Segmentation Toolbox: *AxonSeg*, a software that can segment both axon and myelin from histology images. We will now discuss the axon discrimination, validation results, computation time and results obtained from the three imaging methods assessed, and future perspectives.

4.5.1 Axon detection

AxonSeg was able to detect most of the true axons in OM and SEM images, with a sensitivity of 0.8+ in all three cases tested (see **Table 4.1**). OM presented the best detection results overall, obtaining sensitivities around 0.9 and precisions higher than 0.8 (see "Image Quality and Modalities" section for more details). Better precision was obtained when using a DA classifier for OM and CARS images, for a similar sensitivity. The use of DA classifiers in the CARS sample increases precision by 5%, suggesting that the DA approach could be more robust than the approach without DA for this type of image. It can also be pointed out that CARS results presented lower performance overall when compared to the two other modalities. We believe this could mostly be due to the poor image quality of this particular dataset (sub-optimal fixation, sub-optimal focus, intensity inhomogeneities), the presence of a high number of very small axons (<1μm) in the mouse spinal cord that are difficult to detect due to their size (in comparison, optical microscopy was performed on a cat cervical spinal cord, which has overall larger axons), and the similarity of intensity values between the intracellular compartment and the background, which is inherent to the CARS modality. The benefits of each modality are discussed in section "Image Quality and Modalities".

4.5.2 Discriminant analysis

We believe that the *AxonSeg* software is the first to feature a controllable DA tool allowing more flexibility on the type of contrast/modality, whereas previous software packages have set fixed parameters (e.g. axon shape and size) for detecting axons (More et al., 2011; Begin et al., 2014). Moreover, the user can decide on the ideal sensitivity/specificity depending on the

application. We believe that the best way to exploit the implemented DA model is to build a template classifier by using large dataset of labeled true/false axons for each imaging modality. Therefore, the user could bypass the DA step on the cropped image and simply input one of the template classifiers available.

4.5.3 Segmentation quality

In all three modalities assessed, we observed high Dice coefficients. All Dice medians (50^{th} percentiles) were higher than 0.8. In addition, 10^{th} percentiles were all high, with the lower one being 0.69 (SEM), meaning 90% of axons have a Dice value higher than 0.69. After further analysis of the data, we observed that almost all the Dice coefficients lower than 0.5 come from the smaller axons, which is expected, as the small area (and thus number of pixels) used to calculate the Dice values adds bias to the measures. Note that small differences between the ground truth and segmentation results are expected, as *AxonSeg* produces smooth contours (as described in section "Myelin Segmentation").

Overall, the data obtained shows only small Dice coefficient variations between different imaging techniques, demonstrating the versatility of *AxonSeg*. While other algorithms exist that might produce more robust segmentations (Wang et al., 2011; Liu et al., 2012), we believe that our software is a relevant addition to the existing tools, and its modularity enables other researchers to add more powerful segmentation techniques.

4.5.4 Image quality and modalities

While *AxonSeg* is designed to work with a variety of modalities, further consideration is needed to properly interpret segmentation results in each of the tested modalities. OM images tend to have lower resolution than TEM (Zhao et al., 2010), which can facilitate the segmentation of axons due to the blurring of smaller details in the intracellular compartments (e.g., mitochondria, neurotubules, etc.). However, the blurriness also induces a systematic over-segmentation of the myelin sheath, yielding a downward-bias in the g-ratio (see **Figure 4.5**). In addition, small axons (typically <1μm) are not easily distinguishable at the OM resolution, thereby inducing an upward-bias in axon diameter distribution (Romero et al., 2000) and downward-bias in axon density. SEM images usually have higher resolution and can be noisier, therefore filtering is recommended during pre-processing. Finally, while CARS images produce

exquisite specificity to myelin without the need for staining, images sometimes present some inhomogeneities within the myelin (dark patterns), especially in poorly-fixed samples (Fu et al., 2008; Schie et al., 2015). Moreover, axon and background signal have similar values, which makes it a bit more difficult to detect true axons. Filtering during pre-processing can help improving the detection.

In general, *AxonSeg* like any other segmentation tool strongly depends on the image contrast and resolution. Therefore, the user needs to find the optimal pre-processing parameters for their type of image before proceeding to the next steps. These parameters can then be applied to subsequent histology samples if the image contrast and resolution is kept the same.

4.5.5 Computation time and efficiency

Computation time in *AxonSeg* is reasonably low, needing ~0.26 s processing time per axon. This requires about 8 h for processing 100,000 axons on standard laptop (4 CPUs, 2.9GHz). Note that optimization is ongoing to further reduce processing time (optimal data types, sparse matrices, etc.). *AxonSeg* also includes a feature to automatically partition sub-sections of the full image in order to avoid RAM saturation, as explained in section "Full Image Segmentation".

One of the advantages of *AxonSeg* in terms of implementation efficiency is the *axonlist* structure: every information regarding the objects (axon and myelin segmentation, morphological statistics) is stored in a memory-efficient way, making it easier afterwards to work with the data. In that way, data on the *axonlist* can easily be processed in case the user needs to select specific fibers (e.g. based on position, diameter, and g-ratio).

4.5.6 Distribution of *AxonSeg*

AxonSeg is an open source software, distributed under an MIT licence. It can be downloaded from either Github or MATLAB Central. It is compatible with version R2014a and does not come with compiled binaries, therefore there is no need for an OS cross-compiler. Detailed documentation on how to use *AxonSeg*, including example scripts, sample data and video demos, can be found on Github. Maintenance, optimization and addition of new features are constantly made. Moreover, any researcher is welcome to contribute to the project by adding for instance other detection/segmentation methods and post-processing tools.

4.5.7 Perspectives

Although *AxonSeg* has shown promising results, there is room for improving the quality of the axon detection and segmentation. This is a challenging task because one of the main purposes of *AxonSeg* is to be compatible with as many imaging modalities as possible, in order to be useful for the community at large. This paper focused on presenting the main steps of *AxonSeg* as a proof of concept. Future work aims to explore new segmentation approaches (Markov Random Fields, Level Set), as well as other discrimination strategies, possibly integrating a machine learning module to increase robustness. A complementary approach that is being investigated is an extrapolation of segmentation metrics, robust to false negatives and false positives.

Among possible applications of *AxonSeg* is the validation of quantitative MRI metrics from large-scale histology data, as recently demonstrated by our group (Stikov et al., 2015a; b; Duval et al., 2016).

4.6 Example use case

In this section, we will present an example of workflow and some guidelines to segment an optical microscopy sample. This dataset is available when downloading the *AxonSeg* toolbox, under the "data" folder. Note that there is also a line-by-line tutorial available that shows more examples (*as_tutorial.m*).

4.6.1 Segmentation

The *AxonSeg* GUI can be launched using the following Matlab command line:

```
SegmentationGUI test_image_OM.tif;
```

The pixel size for this histology sample is 0.25 µm. This value can be entered in the corresponding field. This value is important as it will be subsequently used for morphometric statistics.

In this example, axons are bright and myelin is dark. Therefore, the "Invert colour" option needs to be checked. Histogram equalization is recommended in order to improve the axon-to-myelin contrast.

Due to the large size of the image (2000×2000 pixels), the selection of a region of interest is necessary. This can be done by drawing a small rectangle (typically 50-100 µm large) on the image using the "Crop Image" button. A preview window will zoom in this rectangle for the following steps. To go to the next step, click on "Go to next step". To reset parameters and zoom, click on "Reset Step 0".

In the axon segmentation step (step 1), two "ExtendedMin" sliders are available. Two different thresholds can thus be used for the extended minima algorithm if needed. Smaller threshold values do not segment correctly the whole axonal region for bigger axons but are more efficient in order to segment the smaller ones due to the smaller contrast axon/myelin. On the other hand, higher thresholds are useful for segmentation of bigger axons, but tend to merge smaller axons together (if close to each other). Once the thresholds are set, click on "Go to next step" to merge the segmentation (logical OR) from the two sliders. Results are displayed in the preview window.

In the axon discrimination step (step 2), three sliders control different criteria ("Minimal Size", "Solidity", and "Ellipticity") in order to eliminate most of the false positives. Special care is required to avoid the elimination of true positive. Click on "Go to next step".

In the post-processing & myelin segmentation step (step 3), the remaining false positives can be removed using the "Remove" button. Use left click to select axons to remove and right click to validate. Knowing these false positive axons, the discriminant analysis can be applied to improve the specificity of the axon detection. "Quadratic" discriminant analysis usually gives better results in our tests (results not shown). Finally, the "Myelin Segmentation" or "Segment full image (uncropped)" buttons perform the myelin segmentation of the cropped region and full image, respectively.

The *SegParameters.mat* structure contains all segmentation parameters after the myelin segmentation is done on the cropped image. For instance, values that can applied for the segmentation of this sample are:

```
PixelSize=0.25; invertColour=true; HistEq=true; Deconv=false; Smoothing=false;
ExtendedMin1=47.3334; ExtendeMin2=100; minSize=6.4730; Solidity=0.8400;
Ellipticity=0.2100.
```

Any other similar image (same resolution and contrast) can now be segmented using the following command:

```
as_Segmentation_full_image(‘test_image_OM.tif’,’SegParameters.mat’);
```

4.6.2 Visualization

Segmentation results can be displayed in various ways. For example, as myelin sheaths color-coded for g-ratio:

```
load(‘axonlist.mat’);
bw_axonseg=as_display_label(axonlist,size(img),'gRatio','myelin');
display=sc(sc(bw_axonseg,'hot')+sc(img));
imshow(display);
```

Alternatively, both axon and myelin segmentations can be overlaid on the image, color-coded for axon diameter:

```
bw_axonseg_1=as_display_label(axonlist,size(img),'axonEquivDiameter','axon');
bw_axonseg_2=as_display_label(axonlist,size(img),'axonEquivDiameter','myelin');
display=sc(sc(bw_axonseg_1,'hot')+sc(bw_axonseg_2,'hot')+sc(img));
imshow(display);
```

4.6.3 Statistics

Statistics can be computed using the axonlist structure. Below is an example to compute the axon diameter distribution:

```
load(‘axonlist.mat’);
axon_diameters=cat(1,axonlist.axonEquivDiameter);
figure; hist(axon_diameters,50);
```

It is also possible to extract statistics from specific regions of interest (ROI) using a provided RGB mask. Registration between the mask and the segmentation can be performed by running the following code:

```
[mask_reg_labeled, P_color]=as_reg_mask(mask,img);
```

Next, axon IDs belonging to each ROI of the mask are obtained with the following function:

```
indexes=as_stats_mask_labeled(axonlist, mask_reg_labeled);
```

Statistics for each ROI can be displayed as follows:

```
as_stats_barplot(axonlist,indexes,P_color);
```

4.7 References

Alizadeh, A., Dyck, S.M., and Karimi-Abdolrezaee, S. (2015). Myelin damage and repair in pathologic CNS: challenges and prospects. Front Mol Neurosci 8(1662-5099 (Electronic)), 35. doi: 10.3389/fnmol.2015.00035.

Baker, D., and Amor, S. (2014). Experimental autoimmune encephalomyelitis is a good model of multiple sclerosis if used wisely. Multiple Sclerosis and Related Disorders 3(5), 555-564. doi: 10.1016/j.msard.2014.05.002.

Begin, S., Dupont-Therrien, O., Belanger, E., Daradich, A., Laffray, S., De Koninck, Y., et al. (2014). Automated method for the segmentation and morphometry of nerve fibers in large-scale CARS images of spinal cord tissue. Biomed Opt Express 5(12), 4145-4161. doi: 10.1364/boe.5.004145.

Ben-Nun, A., Kaushansky, N., Kawakami, N., Krishnamoorthy, G., Berer, K., Liblau, R., et al. (2014). From classic to spontaneous and humanized models of multiple sclerosis: impact on understanding pathogenesis and drug development. J Autoimmun 54(1095-9157 (Electronic)), 33-50. doi: 10.1016/j.jaut.2014.06.004.

Berthold Ch Fau - Nilsson, I., Nilsson I Fau - Rydmark, M., and Rydmark, M. (1983). Axon diameter and myelin sheath thickness in nerve fibres of the ventral spinal root of the seventh lumbar nerve of the adult and developing cat. (0021-8782 (Print)). doi: D - NLM: PMC1171896 EDAT- 1983/05/01 MHDA- 1983/05/01 00:01 CRDT- 1983/05/01 00:00 PST - ppublish.

Chomiak, T., and Hu, B. (2009). What is the optimal value of the g-ratio for myelinated fibers in the rat CNS? A theoretical approach. PLoS One 4(11), e7754. doi: 10.1371/journal.pone.0007754.

Cuisenaire, O., Romero, E., Veraart, C., and Macq, B.M.M. (Year). "Automatic segmentation and measurement of axons in microscopic images"), 920-929.

Dula, A.N., Gochberg, D.F., Valentine, H.L., Valentine, W.M., and Does, M.D. (2010). Multiexponential T2, magnetization transfer, and quantitative histology in white matter tracts of rat spinal cord. Magn Reson Med 63(4), 902-909. doi: 10.1002/mrm.22267.

Duval, T., Perraud, B., Vuong, M.-T., Lopez Rios, N., Stikov, N., and Cohen-Adad, J. (2016). Validation of quantitative MRI metrics using full slice histology with automatic axon segmentation. Proceedings of the 24th Annual Meeting of ISMRM, Singapore 0928.

Fu, Y., Talavage, T.M., and Cheng, J.X. (2008). New imaging techniques in the diagnosis of multiple sclerosis. Expert Opin Med Diagn 2(9), 1055-1065. doi: 10.1517/17530050802361161.

Lassmann, H. (2014). Mechanisms of white matter damage in multiple sclerosis. Glia 62(1098-1136 (Electronic)), 1816-1830. doi: 10.1002/glia.22597.

Levy, S., Benhamou, M., Naaman, C., Rainville, P., Callot, V., and Cohen-Adad, J. (2015). White matter atlas of the human spinal cord with estimation of partial volume effect. Neuroimage 119(1095-9572 (Electronic)), 262-271. doi: 10.1016/j.neuroimage.2015.06.040.

Liewald, D., Miller, R., Logothetis, N., Wagner, H.J., and Schuz, A. (2014). Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque. Biol Cybern 108(5), 541-557. doi: 10.1007/s00422-014-0626-2.

Liu, T., Jurrus, E., Seyedhosseini, M., Ellisman, M., and Tasdizen, T. (2012). Watershed Merge Tree Classification for Electron Microscopy Image Segmentation. Proc IAPR Int Conf Pattern Recogn 2012, 133-137.

More, H.L., Chen, J., Gibson, E., Donelan, J.M., and Beg, M.F. (2011). A semi-automated method for identifying and measuring myelinated nerve fibers in scanning electron microscope images. J Neurosci Methods 201(1), 149-158. doi: 10.1016/j.jneumeth.2011.07.026.

Papastefanaki, F., and Matsas, R. (2015). From demyelination to remyelination: the road toward therapies for spinal cord injury. Glia 63(7), 1101-1125. doi: 10.1002/glia.22809.

Romero, E., Cuisenaire, O., Denef, J.F., Delbeke, J., Macq, B., and Veraart, C. (2000). Automatic morphometry of nerve histological sections. J Neurosci Methods 97(2), 111-122.

Schie, I.W., Krafft, C., and Popp, J. (2015). Applications of coherent Raman scattering microscopies to clinical and biological studies. Analyst 140(12), 3897-3909. doi: 10.1039/c5an00178a.

Seidl, A.H. (2014). Regulation of conduction time along axons. Neuroscience 276(1873-7544 (Electronic)), 126-134. doi: D - NLM: NIHMS501663.

Stikov, N., Campbell, J.S., Stroh, T., Lavelee, M., Frey, S., Novek, J., et al. (2015a). In vivo histology of the myelin g-ratio with magnetic resonance imaging. Neuroimage 118, 397-405. doi: 10.1016/j.neuroimage.2015.05.023.

Stikov, N., Campbell, J.S., Stroh, T., Lavelee, M., Frey, S., Novek, J., et al. (2015b). Quantitative analysis of the myelin g-ratio from electron microscopy images of the macaque corpus callosum. Data Brief 4, 368-373. doi: 10.1016/j.dib.2015.05.019.

Vincent, L. (1998). Minimal path algorithms for the robust detection of linear features in gray images. Computational Imaging and Vision 12, 331-338.

Wang, Y.Y., Sun Yn Fau - Lin, C.-C.K., Lin Cc Fau - Ju, M.-S., and Ju, M.S. (2011). Segmentation of nerve fibers using multi-level gradient watershed and fuzzy systems. (1873-2860 (Electronic)).

Zhao, X., Pan, Z., Wu, J., Zhou, G., and Zeng, Y. (2010). Automatic identification and morphometry of optic nerve fibers in electron microscopy images. Comput Med Imaging Graph 34(3), 179-184. doi: 10.1016/j.compmedimag.2009.08.009.

Zoupi, L., Savvaki, M., and Karagogeos, D. (2011). Axons and myelinating glia: An intimate contact. IUBMB Life 63(9), 730-735. doi: 10.1002/iub.513.

4.8 Acknowledgments

The authors thank Dr. Serge Rossignol for providing the cat spinal cord, the CM2 at École Polytechnique for providing access to the SEM system, and Ariane Saliani and Blanche Perraud for helping with the acquisition of microscopic data.

4.9 Author Contributions

AZ wrote the paper. AZ and TD designed and developed the software. AG and DC contributed to the code. NS and JCA supervised this work and wrote the paper.

CHAPTER 5 ARTICLE 2: AXONDEEPSEG: AUTOMATIC AXON AND MYELIN SEGMENTATION FROM MICROSCOPY DATA USING CONVOLUTIONAL NEURAL NETWORKS

This paper presents the *AxonDeepSeg* framework designed for axon and myelin segmentation of microscopy images. The segmentation pipeline is based on the design, implementation, training and testing of a convolutional neural network for semantic 3-class segmentation from microscopy images. The network is used to train segmentation models for SEM and TEM modalities. Validation of the models is done on rat and human spinal cord samples (SEM) and mouse and macaque corpus callosum samples (TEM). As a proof of concept, segmentation of a full rat spinal cord is computed and extracted morphological metrics from white matter tracks are compared against the literature.

My **contributions** include the participation in the design and implementation of the framework (collection and management of datasets, experiments on data augmentation and dropout strategies, network design, implementation of the testing and morphometrics extraction modules), as well as the generation of the results, writing of the paper and the generation of the figures. This paper was published in *Nature Scientific Reports*.

doi: 10.3389/fninf.2016.00037.

Authors: Aldo Zaimi[+,1], Maxime Wabartha[+,1,2], Victor Herman[1,2], Pierre-Louis Antonsanti[1,3], Christian S. Perone[1], Julien Cohen-Adad[1,4].

Affiliations:

[1]NeuroPoly Lab, Institute of Biomedical Engineering, Polytechnique Montreal, Montreal, QC, Canada;

[2]École Centrale de Lille, Lille, France;

[3]École Centrale de Nantes, Nantes, France;

[4]Functional Neuroimaging Unit, CRIUGM, Université de Montréal, Montreal, QC, Canada.

([+]) These authors contributed equally to this work.

5.1 Abstract

Segmentation of axon and myelin from microscopy images of the nervous system provides useful quantitative information about the tissue microstructure, such as axon density and myelin thickness. This could be used for instance to document cell morphometry across species, or to validate novel non-invasive quantitative magnetic resonance imaging techniques. Most currently-available segmentation algorithms are based on standard image processing and usually require multiple processing steps and/or parameter tuning by the user to adapt to different modalities. Moreover, only a few methods are publicly available. We introduce *AxonDeepSeg*, an open-source software that performs axon and myelin segmentation of microscopic images using deep learning. *AxonDeepSeg* features: (i) a convolutional neural network architecture; (ii) an easy training procedure to generate new models based on manually-labeled data and (iii) two ready-to-use models trained from scanning electron microscopy (SEM) and transmission electron microscopy (TEM). Results show high pixel-wise accuracy across various species: 85% on rat SEM, 81% on human SEM, 95% on mice TEM and 84% on macaque TEM. Segmentation of a full rat spinal cord slice is computed and morphological metrics are extracted and compared against the literature. *AxonDeepSeg* is freely available at https://github.com/neuropoly/axondeepseg.

5.2 Introduction

Neuronal communication is ensured by the transmission of action potentials along white matter axons. For long distance communication, these axons, which are typically 1-10µm in diameter, are surrounded by a myelin sheath whose main role is to facilitate the propagation of the electrical impulses along neuronal fibers and increase the transmission speed[1,2]. Pathologies such as neurodegenerative diseases (e.g., multiple sclerosis) or trauma are associated with myelin degeneration, which can ultimately lead to sensory and motor deficits (e.g., paraplegia)[3,4]. Being able to image axons and myelin sheaths at high resolution would help researchers understand the origins of demyelination and test therapeutic drugs[5,6] and could also be used to validate novel magnetic resonance imaging biomarkers of myelin[7]. High resolution histology is typically done using electron microscopy following osmium staining to obtain myelin contrast. Then, axons and myelin can be analysed on the images to derive metrics such as axon density or myelin thickness.

However, given that 1 mm^2 of white matter can contain over 100,000 axons[8], it is important to obtain a robust and reliable segmentation of individual axons and myelin as automatically as possible.

Several segmentation methods for axon and myelin have been proposed which are based on traditional image processing algorithms including thresholding and morphological operations[9,10], axon shape-based morphological discrimination[11], watershed[12,13], region growing[14], active contours without[15,16] and with discriminant analysis[16]. However, a few limitations can be reported from the previous work: (i) traditional image-based methods are designed to work on specific imaging modalities and often fail if another contrast is used (e.g., optical image instead of electron microscopy); (ii) previous methods are not fully-automatic as they typically require either preprocessing, hand-selected features for axon discrimination and/or postprocessing; (iii) traditional image-based methods do not make full use of the contextual information of the image (i.e., multi-scale representation of axons, average shape of axons, etc.) and (iv) most of the previous methods are not publicly available (to our knowledge, only that from[15,16] are).

In the last five years, deep learning methods have become the state of the art when it comes to computer vision tasks. Convolutional neural networks (CNNs) are particularly suited to image classification[17–20] and semantic segmentation[21]. Cell segmentation is one of the popular application of CNNs[22,23]. The U-Net architecture introduced by Ronneberger and collaborators[24] has inspired many medical segmentation applications, efficiently combining both context and localization of structures of interest. Segmentation of axons and myelin based on deep learning approaches offers significant advantages when compared with traditional image segmentation algorithms: (i) there is no need to hand-select relevant features because the network is able to learn the hidden structural and textural features by itself, (ii) this approach allows to segment both axons and myelin sheaths in two different labels with the same network, without the need of any explicit pre- or post-processing, (iii) the network can be trained for various imaging modalities without significantly changing its architecture and (iv) once trained, the model is relatively fast at the prediction step (only a few seconds) compared to more traditional image processing methods.

Few research groups have applied deep learning for axon and myelin segmentation. Naito and collaborators[25] have implemented a two-step process that first performs clustering segmentation of myelinated nerve fibers in optical microscopic images, and then discriminates between true and false candidates by using a CNN classification network. This group did not exploit the CNN for the segmentation, but only for discrimination. The work from Mesbah and collaborators[26] presented a deep encoder-decoder CNN that can segment both axon and myelin and claimed to achieve up to 82% pixel-wise accuracy. However, the network has been designed specifically for light microscopy images, the implementation is not publicly available and minimal regularization strategies have been employed in order to improve generalization.

We present *AxonDeepSeg*, a deep learning framework for robust and automatic segmentation of both axons and myelin sheaths in myelinated fibers. *AxonDeepSeg* features: (i) a CNN architecture for semantic segmentation of histological images; (ii) two ready-to-use models for the segmentation of scanning electron microscopy (SEM) and transmission electron microscopy (TEM) samples adapted to a variety of species and acquisition parameters; (iii) a well-documented training pipeline to generate models for new imaging modalities and (iv) free and open source code (https://github.com/neuropoly/axondeepseg).

5.3 Methods

5.3.1 Dataset

Microscopy images used in this study were acquired with two different imaging techniques: SEM and TEM. Different acquisition resolutions were used, in order to increase variability and obtain better generalization of the model, with isotropic pixel size resolution ranging from 0.05 to 0.18 μm (SEM) and 0.002 to 0.009 μm (TEM). SEM samples were stained with 2% osmium, embedded in epoxy, polished and imaged with the same SEM system (Jeol 7600F). TEM images were obtained from mice brain samples (splenium), as described in[27]. Additionally, a macaque sample of the corpus callosum was added to the test set. Preparation and imaging procedures are described in[7]. **Table 5.1** lists the samples used for the experiments.

All methods were carried out in accordance with relevant guidelines and regulations. Experimental protocols involving rats were approved by the Montreal Heart Institute committee. Experimental protocols involving the human spinal cord were done at the anatomy laboratory of

the University of Quebec at Trois-Rivieres. The spinal cord donor gave informed consent and procedures were approved by the local ethics committee (SCELERA-15-03-pr01). Similarly, TEM images shared by collaborators were obtained in accordance with the corresponding ethics committees (mice: Institutional Animal Care and Use Committee at the New York University School of Medicine, macaque: Montreal Neurological Institute Animal Care Committee).

5.3.2 Ground truth labeling

The ground truth labeling of SEM samples was created as follows: (i) Myelin sheaths were manually segmented (inner and outer contours) with GIMP (https://www.gimp.org/); (ii) Axon labels were obtained by filling the region enclosed by the inner border of the myelin sheaths; (iii) Small manual corrections were done on the axon and myelin masks (contour refinement, elimination of false positives) when necessary.

The ground truth labeling of TEM samples was created as follows: (i) Myelin was first segmented using intensity thresholding followed by manual correction, then the inner region was filled to generate axon labels. More details can be found about the generation of labels for the macaque[7] and the mice[27].

All ground truth labels were cross-checked by at least two researchers. The final ground truth consists of a single *png* image with values: background=0, myelin=127, axon=255. Example SEM and TEM samples and corresponding ground truth labels are shown in **Figure 5.1**. This figure also illustrates the large variability in terms of image features, especially for the SEM data (contrast, noise, sample preservation, etc.).

Table 5.1: List of datasets used for the experiments. For each sample, the following information is indicated: number of images used, species, tissue type, pixel size, field of view (FOV) and tissue preparation details. For the scanning electron microscopy (SEM) model, training was done on rat spinal cord samples and testing was performed on rat and human spinal cord samples. For the transmission electron microscopy (TEM) model, training was done on mice brain samples and testing was performed on mice and macaque brain samples.

		Number of images	**Species**	**Tissue**	**Pixel size (µm)**	**FOV (µm^2)**	**Tissue preparation (% paraformaldehyde – % glutaraldehyde)**
SEM	Training / validation	1	Rat	Spinal cord (cervical)	0.18	230×166	4% – 2%
		3	Rat	Spinal cord (cervical)	Between 0.05 and 0.17	Between 132×90 and 218×162	4% – 0%
		3	Rat	Spinal cord (cervical)	0.1	Between 74×76 and 77×84	3% – 3%
		1	Rat	Spinal cord (cervical)	0.13	247×234	3% – 3%
		1	Rat	Spinal cord (cervical)	0.1	82×77	3% – 3%
	Testing	1	Rat	Spinal cord (cervical)	0.13	150×97	3% – 3%
		1	Rat	Spinal cord (cervical)	0.07	108×77	3% – 3%
		1	Human	Spinal cord (cervical)	0.13	715×735	4% – 2%
TEM	Training / validation	8 × 17 mice	Mouse	Brain (splenium)	0.002	6×9	2% – 2.5%
	Testing	8 × 3 mice	Mouse	Brain (splenium)	0.002	6×9	2% – 2.5%
		1	Macaque	Brain (corpus callosum)	0.009	27×21	2% – 2%

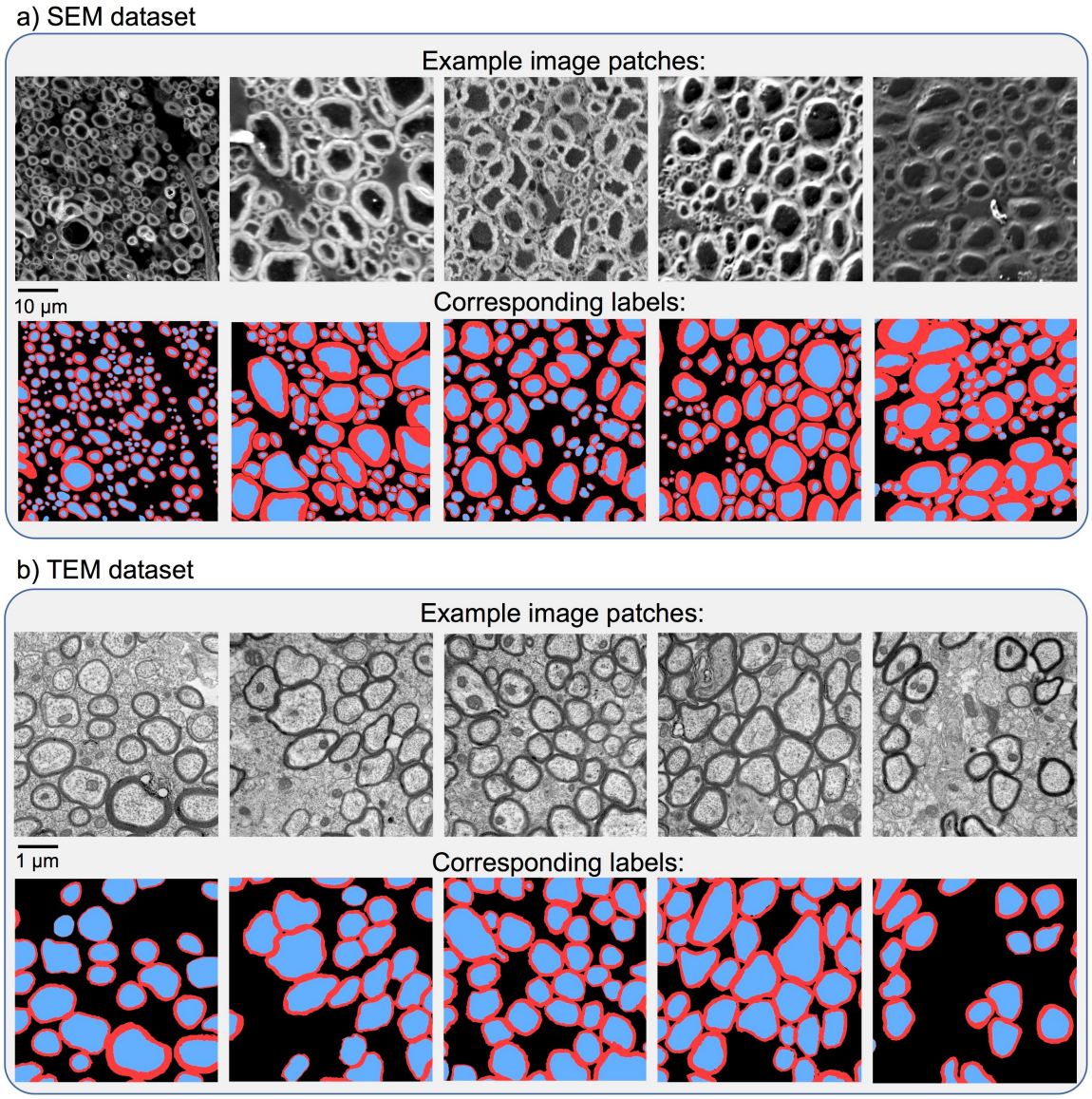

Figure 5.1: Overview of the data and ground truth labels for SEM (a) and TEM (b). Label masks contain 3 classes: axon (in blue in the figure), myelin (red) and background (black). All SEM and TEM samples shown here are cropped to 512×512 pixels. SEM patches have a pixel size of 0.1 µm, while TEM patches have a pixel size of 0.01 µm (see section "Pipeline overview").

5.3.3 Pipeline overview

The pipeline of *AxonDeepSeg* is composed of four steps: data preparation, learning, evaluation and prediction. **Figure 5.2** illustrates each step.

In the data preparation step, raw microscopy images and corresponding axon/myelin labels are resampled to a common resolution space: 0.1 μm per pixel for SEM and 0.01 μm for TEM. These values are based on preliminary results and on the typical resolutions provided by each of these imaging systems. Resampled samples are divided into patches of 512×512 pixels due to memory constraints. This size was chosen to have around 15-75 axons per patch. Traditional pre-processing was applied patch-wise, including standardization and histogram equalization (not shown in figure 2 for clarity). For learning, the patches and corresponding labels were randomly split and then considered either for the training or for the validation sets (training/validation split of approximately 70/30%). For evaluation, full test images were randomly selected.

In the learning step, the training/validation dataset is fed into the network. Once the trained model is obtained, performance is evaluated on the test dataset (evaluation step). Finally, the trained model can be used for inference on new microscopy images (prediction step). The images are resampled to the pixel size of the model, divided into patches of 512×512 pixels, segmented, stitched to the native size, and resampled to the native resolution. Note that bilinear interpolation was used during the resampling steps.

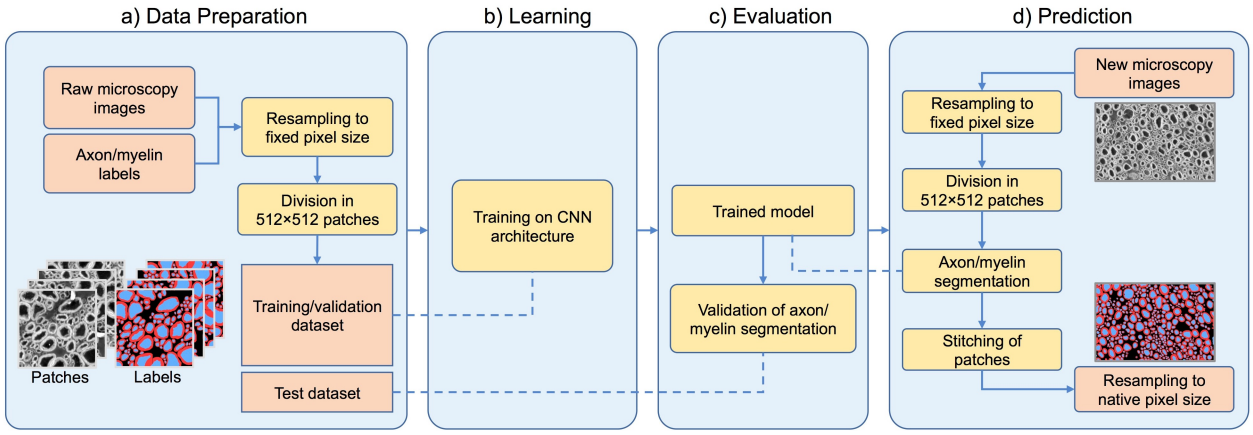

Figure 5.2: Overview of the *AxonDeepSeg* pipeline. During the data preparation step (a), microscopy samples and corresponding ground truth labels are resampled to have a common pixel size (0.1 µm for the SEM model, 0.01 µm for the TEM model), divided into 512×512 patches, and split into training/validation sets. The neural network is trained during the learning step (b) on the training/validation dataset. When the model is trained, performance is assessed on a test dataset (evaluation step (c)). For prediction (d), the new microscopy image to be segmented is first resampled to the working pixel size of the network, divided into 512×512 patches and analysed with the trained model. Segmented output patches are then stitched together and resampled back to the native pixel size.

5.3.4 Architecture of the network

The architecture is inspired by the original U-Net model[24], combining a contracting path with traditional convolutions and then an expanding path with up-convolutions. **Figure 5.3** illustrates the network architecture. The convolutional layers in the first block use 5×5 kernels, while the convolutional layers on remaining blocks use 3×3 kernels. The SEM network has 3 convolutional layers per block, while the TEM network has 2 convolutional layers per block. These decisions were based on preliminary optimizations (see section "Hyperparameter optimization"). In the contracting path, convolutions of stride 2 are computed after the last convolutional layer of each block to reduce the dimensionality of the features. Each strided convolution layer has a corresponding up-convolution layer in the expansion path in order to recover the localization information lost during the contraction path. Up-convolutions were computed by bilinear interpolation followed by a convolution. The merging of the context and localization information is done by concatenating the features from the contracting path with the

corresponding ones in the expansion path. The number of features (channels) is doubled after each block, starting from 16, and then decreased at the same rate during the expansion path. All activation functions in the convolutional layers are rectified linear units (ReLU[28]). The last layer before the prediction is a softmax activation with 3 classes (axon, myelin and background). The SEM and TEM networks have a total of 1,953,219 and 1,552,387 trainable parameters, respectively.

5.3.5 Data augmentation strategy

A data augmentation strategy was used on the input patches in order to reduce overfitting and improve generalization[17,20,24]. The strategy includes random shifting, rotation, rescaling, flipping, blurring and elastic deformation[29]. **Table 5.2** summarizes the data augmentation strategy and the corresponding parameters.

Table 5.2: Data augmentation strategy used in *AxonDeepSeg*. Shifting, rotation, rescaling, flipping, blurring and elastic deformation were applied to training patches in order to reduce overfitting and increase variability.

Data augmentation strategy	Description
Shifting	Random horizontal and vertical shifting between 0 and 10% of the patch size, sampled from a uniform distribution.
Rotation	Random rotation, angle between 5 and 89 degrees, sampled from a uniform distribution.
Rescaling	Random rescaling of a randomly sampled factor between 1/1.2 and 1.2
Flipping	Random flipping: vertical flipping or horizontal flipping.
Blurring	Random blurring: gaussian blur with the standard deviation of the gaussian kernel being uniformly sampled between 0 and 4.
Elastic deformation	Random elastic deformation with uniformly sampled deformation coefficient α=[1-8] and fixed standard deviation σ=4.

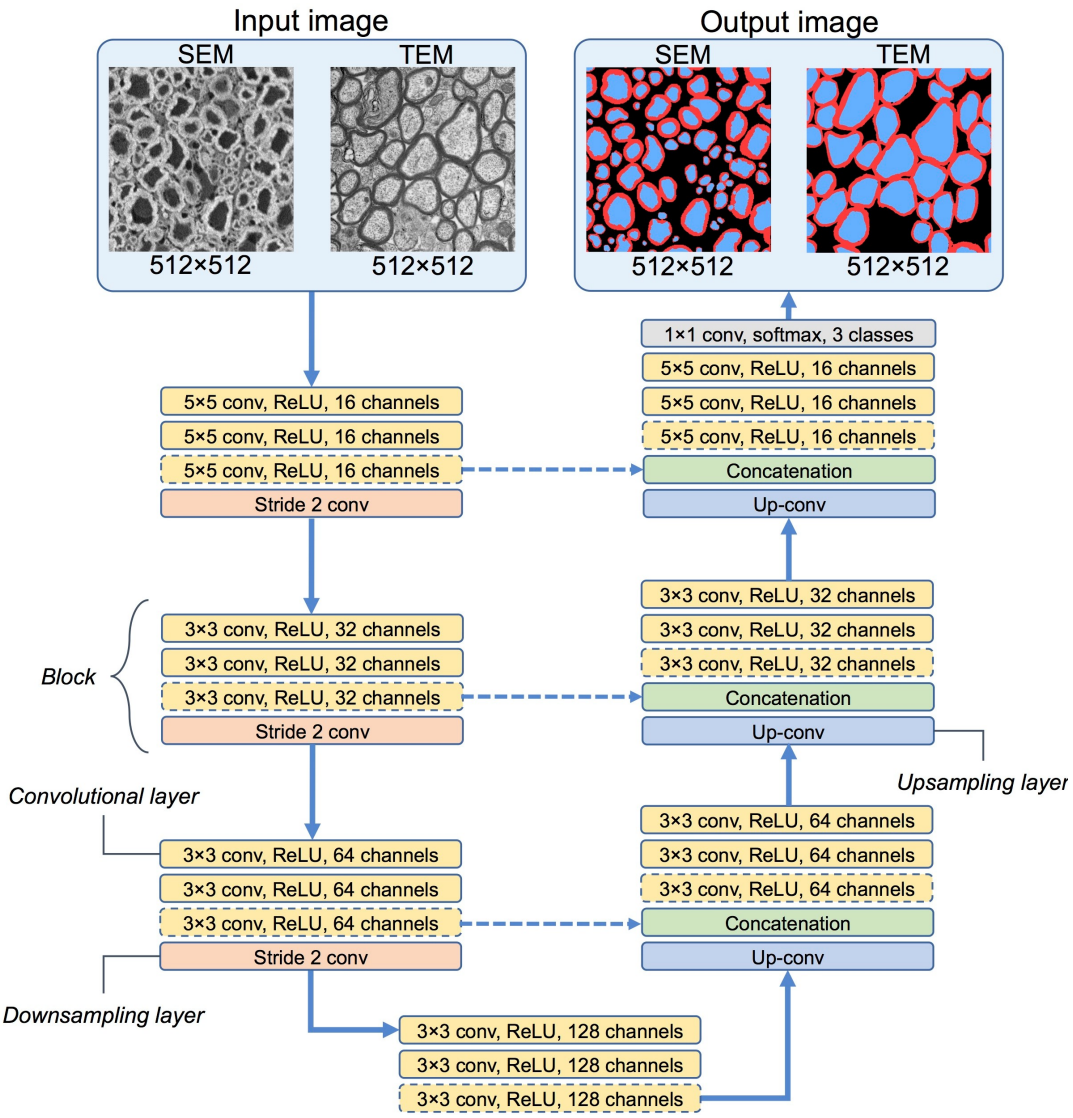

Figure 5.3: Architecture of the convolutional neural networks designed for the segmentation of SEM and TEM images. For the SEM model, 3 convolutional layers are used at each block, while only 2 convolutional layers are used for the TEM model. Convolutional layers in dashed lines are removed for the TEM model. All activation functions used are rectified linear units (ReLU). Strided convolutions are used to downsample the features during the contraction path (left), while up-convolutions are used to recover the localization during the expansion path (right). Features of the contraction path are merged with features of the expansion path to combine localization and context (illustrated by the concatenation step). The pixel-wise classification is done by a 3-class softmax.

5.3.6 Training procedure

For the training phase, we used a starting learning rate of 0.001 on which we applied a polynomial decay[30] with a power of 0.9. The decay length was 200 epochs, after which the training stopped. We shuffled the samples list at the beginning of each epoch and used a batch size of 8 patches of 512×512 pixels. We have also implemented batch normalization[31] before each activation. The momentum was exponentially decayed from 0.7 to 0.9. This was done to enable a quicker convergence at the beginning of the training by keeping a few samples for the batch normalization, while ensuring a stable training at the later epochs. A dropout[32] rate of 0.25 is used in the convolutional layers to reduce the risk of overfitting and improve generalization. The network was trained with the Adam optimizer[33]. We minimized a spatially-weighted multi-class cross-entropy loss. The spatial weights ratios used to correct the class imbalance were respectively 1.1, 1.0 and 1.3 for background, myelin and axon. Those weights were chosen after hyperparameter optimization. The training phase took 86 minutes on an NVIDIA P100 GPU.

5.3.7 Inference procedure

During the inference step, we split the original images into patches of size 512×512 pixels. To overcome border issues (i.e. partial axons at edges not being properly identified as axons), the output segmentation mask is cropped around a smaller patch. Thus, patches overlap by *d* pixels to cover the entire image, as illustrated in **Figure 5.4**. Based on preliminary optimizations, the default value *d* was set to 25.

5.3.8 Hyperparameter optimization

We used different grid searches in order to set the value of the hyperparameters with respect to the accuracy and error on the validation set. The following architecture parameters were optimized at the same time: number of layers, number of filters and convolutional kernel size. The starting learning rate and the batch normalization momentum were also optimized jointly using a grid search, as they both have an effect on the time the model takes to converge and the stability of the validation metrics (based on our experiments). We then jointly optimized the batch normalization momentum and the decay period of the momentum.

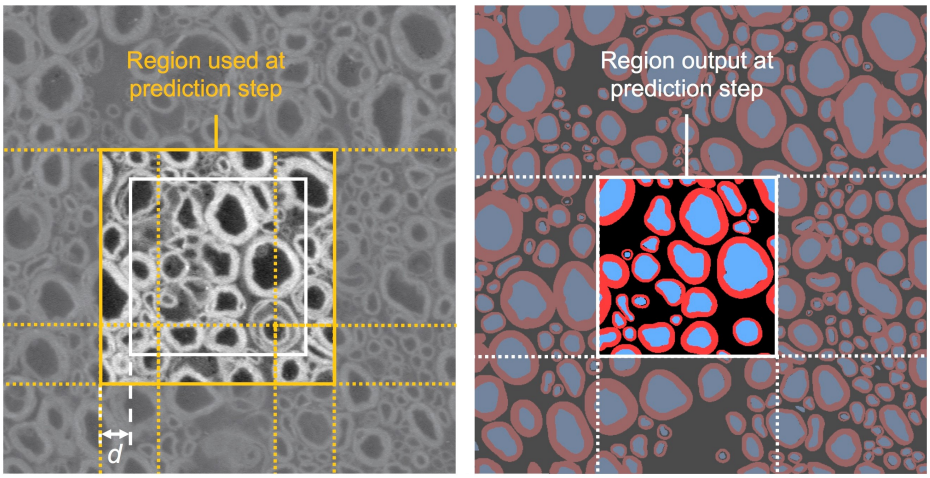

Figure 5.4: Overlapping procedure during inference. To avoid border effects during prediction, inference is run on the orange square, but only the white square is output. The algorithm iterates by shifting the inference window by the size of the white square. The overlap default value *d* was set to 25.

5.3.9 Evaluation method

For testing, the following metrics were computed: the Dice values (axon and myelin) and the pixel-wise accuracy to assess the quality of the segmentation, and the sensitivity and precision to assess the capability to detect true axonal fibers and avoid false axonal fibers.

5.3.9.1 Segmentation metrics

To assess the quality of the segmentation we used the Dice coefficient. For two binary images *A* and *B*, the Dice coefficient is defined as:

$$Dice = \frac{2(A \cap B)}{|A| + |B|} \tag{5.1}$$

where $A \cap B$ is the intersection between the two images (i.e. number of pixels that are true in both images), $|A|$ is the number of pixels that are true in image *A*, and $|B|$ is the number of pixels that are true in image *B*. The Dice coefficient is computed separately for axon and myelin segmentations, between the prediction and the ground truth masks.

Furthermore, the pixel-wise accuracy is evaluated in order to get a combined assessment of axon-myelin segmentation. The pixel-wise accuracy is computed as the ratio between correctly classified pixels (i.e. axon pixel classified as axon, myelin pixel classified as myelin, background pixel classified as background) and the total number of pixels in the test sample.

5.3.9.2 Detection metrics

To assess the performance of myelinated fiber detection, we computed the sensitivity and precision based on axon objects, using the positions of the centroids. Knowing the number of true positives (TP, axons present in both the prediction and the ground truth mask), false positives (FP, axons present in the prediction, but absent in the ground truth mask) and false negatives (FN, axons present in the ground truth mask, but absent in the prediction), we can compute the sensitivity (true positive rate) and the precision (positive predictive value) with the following equations:

$$TPR = \frac{TP}{TP + FN} \tag{5.2}$$

$$PPV = \frac{TP}{TP + FP} \tag{5.3}$$

5.3.10 Data availability

A part of the datasets generated during and/or analysed during the current study are available in the *White Matter Microscopy Database* repository (https://osf.io/yp4qg/). The remaining datasets are available from the corresponding author on reasonable request.

5.4 Results

5.4.1 Segmentation

Segmentation was evaluated on SEM (rat and human spinal cords) and TEM (mouse splenium and macaque corpus callosum) samples. Segmentation and ground truth masks for both axons and myelin sheaths are displayed on **Figure 5.5**. **Table 5.3** lists validation metrics computed on the segmentation outputs: axon Dice, myelin Dice, pixel-wise accuracy, sensitivity

and precision. The SEM model trained on rat microscopy was able to achieve a pixel-wise accuracy between 85% and 88% on the rat test samples, while the pixel-wise accuracy on human test sample was 81%. The TEM model trained on mice microscopy achieved a pixel-wise accuracy of 95% on mice samples and a pixel-wise accuracy of 84% on macaque samples.

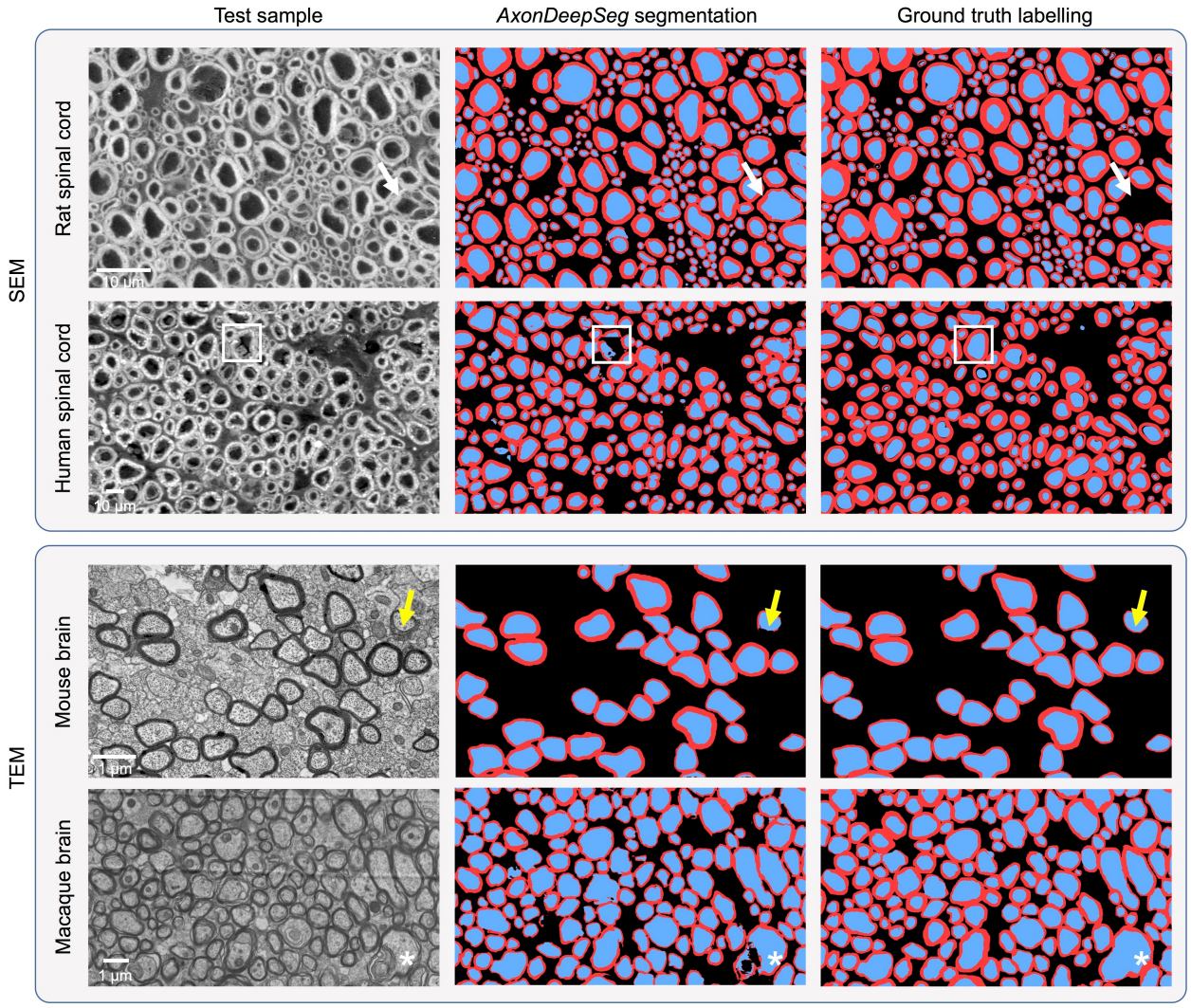

Figure 5.5: Example of segmentation results on SEM and TEM images on a variety of species. The corresponding ground truth segmentation is shown on the right. Overall, the agreement is good. A few discrepancies are noticeable, notably caused by ambiguous/untypical myelin structure (white arrows and white asterisks), inhomogeneous myelin thickness (yellow arrows) and untypical axon intensity (white squares). Some of these discrepancies could potentially be solved using post-processing methods.

Table 5.3: Summary of performance metrics on test samples, for both SEM and TEM models. The SEM model was trained on rat spinal cord samples, and evaluated on rat and human spinal cord samples, while the TEM model was trained on mice brain samples and evaluated on mice and macaque brain samples. For each sample, axon Dice, myelin Dice, pixel-wise accuracy, sensitivity and precision were computed. Axon and myelin Dice measure the similarity between the axon/myelin segmentation masks and the ground truth. Pixel-wise accuracy is a measure of the ratio of correctly classified pixels. Sensitivity and precision values are an indication of the capability to detect true axonal fibers and to avoid segmentation of false axonal fibers. Note that for the mice, 24 samples of the same size were used: performance metrics shown are means between all samples.

Modality	**Model**	**Test sample(s)**	**Axon Dice similarity**	**Myelin Dice similarity**	**Pixel-wise accuracy**	**Sensitivity**	**Precision**
SEM	Trained on rat samples	Rat 1	0.9089	0.8193	0.8510	0.9699	0.8468
		Rat 2	0.9244	0.8389	0.8822	0.9876	0.7987
		Human	0.8089	0.7629	0.8114	0.9300	0.7306
TEM	Trained on mice samples	Mice	0.9493	0.8552	0.9451	0.9597	0.9647
		Macaque	0.9069	0.7519	0.8438	0.9429	0.8129

To demonstrate the utility of *AxonDeepSeg* for large scale microscopy, segmentation of axon/myelin was performed on a full rat spinal cord SEM (cervical level). Processing time was 5 hours on a Mac laptop (2.9 GHz). Segmentation masks (axons in red, myelin sheaths in blue) are displayed on **Figure 5.6**, along with a zoomed window of a small region for better visualization.

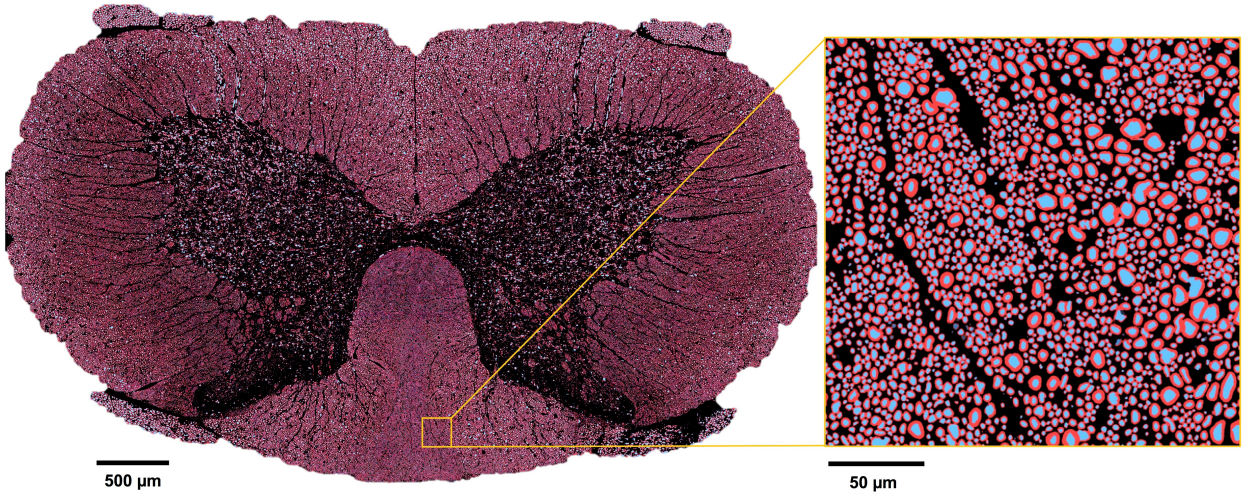

Figure 5.6: Full slice of rat spinal cord showing segmented axons (blue) and myelin sheaths (red). The zoomed panel illustrates the segmentation performance and sensitivity to fiber size: the left half of the panel contains smaller axons (mean diameter around 1.75 µm) while the right half contains larger axons (mean diameter around 2.5 µm).

5.4.2 Morphometrics extraction

As a proof-of-concept, morphometric statistics were extracted from a full spinal cord of rat using *AxonSeg*[16]. The segmented rat spinal cord shown in **Figure 5.6** was downsampled to 50×50 µm^2 in order to generate maps of density (e.g., axon and myelin density). The following aggregate metrics were computed:

- Axon diameter mean and standard deviation: arithmetic mean and standard deviation of the distribution of equivalent axon diameters (computed for each axon object as $\sqrt{(4\times Area/\pi)}$);
- Axon density: number of axons per mm^2;
- Axon volume fraction (AVF): ratio between area of axons and total area of the region;
- Myelin volume fraction (MVF): ratio between area of myelin and total area of the region;
- G-ratio: ratio between axon diameter and myelinated fiber (axon + myelin) diameter, which can be estimated with the following formula[7]: $\sqrt{(1/(1+MVF/AVF))}$.

A binary mask was used to only keep white matter pixels. Results are displayed in **Figure 5.7**. Obtained metrics were compared with references of the white matter tracts of the rat spinal

cord[34–36]. The distribution maps are in good agreement with known anatomy. In the corticospinal tract (tract #12 of the reference), we observe smaller axon diameters (around 1 µm), very high axon density (around 200,000 axons per mm^2) and g-ratio values around 0.6. Larger axons are found close to the spinal cord periphery. See Discussion for more comparisons with the literature.

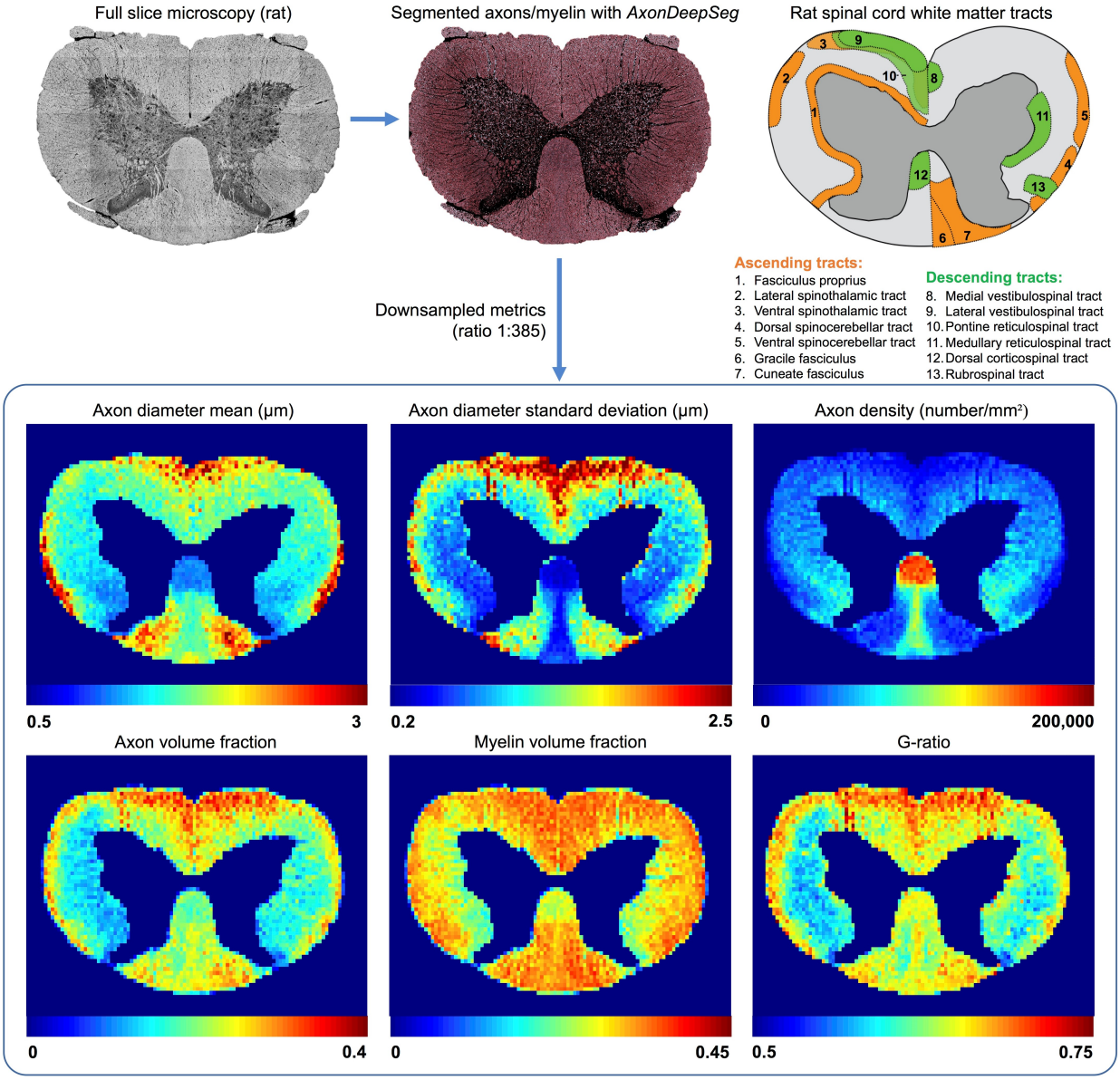

Figure 5.7: Distribution maps of axon diameter mean and standard deviation, axon density, axon volume fraction, myelin volume fraction and g-ratio in a full rat spinal cord slice (cervical level). The SEM slice was segmented with *AxonDeepSeg*. The aggregate metrics of the white matter were generated by downsampling the axon/myelin segmentation masks to a 50×50 µm^2

resolution. A schematic diagram of the main ascending and descending tracts of the white matter in the rat spinal cord based on the literature[34–36] is provided as reference.

5.5 Discussion

This paper introduced *AxonDeepSeg*, a software framework to segment axon and myelin from microscopy data using deep learning. We showed that *AxonDeepSeg* can segment axon and myelin of SEM and TEM samples of various species with high accuracy. Moreover, *AxonDeepSeg* can serve as a tool to document nerve fiber morphometry, as demonstrated by the extraction of metrics from a full rat spinal cord slice.

5.5.1 Trained models

We propose a SEM model trained with a resolution of 0.1 μm per pixel, and a TEM model trained with a resolution of 0.01 μm per pixel. At inference, test image is resampled to meet the target resolution of the model. Other training set compositions were explored, with model trained on both SEM and TEM data in order to achieve better generalization. However, a few limitations arose: (i) SEM and TEM images exhibit very different resolution ranges, requiring large resampling factors to find a common resolution space; (ii) SEM and TEM modalities capture different microstructure/textures of the tissue (for instance, TEM microscopy can capture subcellular microstructure details of the axon); (iii) preliminary results of model simultaneously trained on SEM and TEM led to lower performance when compared to modality-specific models.

5.5.2 Performance metrics

In all test sets, sensitivity was high (>93%, see **Table 5.3**), indicating good capability to detect true positive axons. Lower performance metrics obtained in the human SEM sample are expected, as the human sample used exhibits different contrast/quality/noise properties when compared to the rat training set. Note that myelin sheaths of the macaque TEM sample are slightly underestimated when compared to the ground truth segmentation. In both models and all test samples, computed myelin Dice was lower than axon Dice. This could be explained by the fact that myelin objects have two interfaces: boundary ambiguity between myelin and axon, and boundary ambiguity between myelin and background. Therefore, the myelin Dice is affected by

two types of myelin misclassifications: myelin pixel classified as axon or myelin pixel classified as background.

Overall, these results suggest that the trained SEM and TEM models are robust to a variety of species and contrast changes and can generalize well, given that the lowest pixel-wise accuracy observed was 81% (see **Table 5.3**). Similar work done on optical microscopy data[26] have achieved a maximal pixel-wise accuracy of 82%. As pointed out in **Figure 5.5**, most pixel misclassifications are due to ambiguous/untypical axon and/or myelin structure or intensity distribution. Note that these discrepancies could possibly be solved by implementing post-processing methods based on mathematical morphology or conditional random fields.

5.5.3 Morphometrics extraction

Morphological metrics were extracted from a full rat spinal cord slice at the cervical level (see **Figure 5.7**). The metrics resulting from the segmentation are overall consistent with the known anatomy. The ventral spinothalamic tract (#3 in tract reference of **Figure 5.7**) contains the largest axons[34,36], while higher density and smaller axons are observed in the corticospinal tract (#12 in tract reference)[35,36]. Furthermore, the spinocerebellar tracts (#4 and #5 in tract reference) are mostly composed of large diameter fibers[34]. We also observe that axons in the cuneate fasciculus (#7 in tract reference) are larger than those found in the gracile fasciculus (#6 in tract reference), which is also in agreement with the literature[37]. G-ratio ranges between 0.5 and 0.75, which is in agreement with other rat microstructure studies[38]. Overall, concordance of metrics obtained with literature shows that *AxonDeepSeg* can serve as a tool to document distribution and size of myelinated fibers in microscopy samples.

5.5.4 Software

AxonDeepSeg is coded in Python and based on the *TensorFlow* deep learning framework. It can currently run on Linux and Mac OS X systems. Segmentation inference can be done on standard CPU computers at reasonable computational time. The code is available as open source in GitHub (https://github.com/neuropoly/axondeepseg) and an intuitive documentation is provided (https://neuropoly.github.io/axondeepseg/). A Binder link and a simple Jupyter notebook are available for getting started with *AxonDeepSeg*.

5.5.5 Future perspectives

The use of ensemble techniques, which consist of combining multiple neural network models, can potentially increase performance metrics. However, its drawback is that it increases computational time at inference. Another possible approach is to use transfer learning[39] in order to obtain better generalization in new imaging modalities even when having a small training set. A partially trained model can be used as starting point for the training of another model of different modality. Note that *AxonDeepSeg* has been trained and tested on healthy tissues. It would be interesting to assess its performance on demyelinated microscopy samples, in which myelin sheaths might present smaller thickness and different morphology.

Even though current models perform well, our long-term goal is to continuously improve these models by adding more training data from collaborators in order to improve generalization. Another objective is to build segmentation models for other modalities, such as optical microscopy and Coherent Anti-Stokes Raman spectroscopy (CARS). This vision is supported by the recent initiative of creating a White Matter Microscopy Database[40], which provides to the community an open access microscopy data and associated labeled ground truth. We encourage people to share their data for fostering the development of performant segmentation methods.

5.6 References

1. Zoupi, L., Savvaki, M. & Karagogeos, D. Axons and myelinating glia: An intimate contact. IUBMB Life 63, 730–735 (2011).

2. Seidl, A. H. Regulation of conduction time along axons. Neuroscience 276, 126–134 (2014).

3. Lassmann, H. Mechanisms of white matter damage in multiple sclerosis. Glia 62, 1816–1830 (2014).

4. Papastefanaki, F. & Matsas, R. From demyelination to remyelination: the road toward therapies for spinal cord injury. Glia 63, 1101–1125 (2015).

5. Sachs, H. H., Bercury, K. K., Popescu, D. C., Narayanan, S. P. & Macklin, W. B. A new model of cuprizone-mediated demyelination/remyelination. ASN Neuro 6, (2014).

6. Pfeifenbring, S., Nessler, S., Wegner, C., Stadelmann, C. & Brück, W. Remyelination After Cuprizone-Induced Demyelination Is Accelerated in Juvenile Mice. J. Neuropathol. Exp. Neurol. 74, 756–766 (2015).

7. Stikov, N. et al. In vivo histology of the myelin g-ratio with magnetic resonance imaging. Neuroimage 118, 397–405 (2015).

8. Saliani, A. et al. Axon and Myelin Morphology in Animal and Human Spinal Cord. Front. Neuroanat. 11, 129 (2017).

9. Romero, E. et al. Automatic morphometry of nerve histological sections. J. Neurosci. Methods 97, 111–122 (2000).

10. Cuisenaire, O., Romero, E., Veraart, C. & Macq, B. M. M. Automatic segmentation and measurement of axons in microscopic images. in Medical Imaging 1999: Image Processing 3661, 920–930 (International Society for Optics and Photonics, 1999).

11. More, H. L., Chen, J., Gibson, E., Donelan, J. M. & Beg, M. F. A semi-automated method for identifying and measuring myelinated nerve fibers in scanning electron microscope images. J. Neurosci. Methods 201, 149–158 (2011).

12. Liu, T., Jurrus, E., Seyedhosseini, M., Ellisman, M. & Tasdizen, T. Watershed Merge Tree Classification for Electron Microscopy Image Segmentation. Proc. IAPR Int. Conf. Pattern Recogn. 2012, 133–137 (2012).

13. Wang, Y.-Y., Sun, Y.-N., Lin, C.-C. K. & Ju, M.-S. Segmentation of nerve fibers using multi-level gradient watershed and fuzzy systems. Artif. Intell. Med. 54, 189–200 (2012).

14. Zhao, X., Pan, Z., Wu, J., Zhou, G. & Zeng, Y. Automatic identification and morphometry of optic nerve fibers in electron microscopy images. Comput. Med. Imaging Graph. 34, 179–184 (2010).

15. Bégin, S. et al. Automated method for the segmentation and morphometry of nerve fibers in large-scale CARS images of spinal cord tissue. Biomed. Opt. Express 5, 4145–4161 (2014).

16. Zaimi, A. et al. AxonSeg: Open Source Software for Axon and Myelin Segmentation and Morphometric Analysis. Front. Neuroinform. 10, 37 (2016).

17. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. in Advances in Neural Information Processing Systems 25 (eds. Pereira, F., Burges, C. J. C., Bottou, L. & Weinberger, K. Q.) 1097–1105 (Curran Associates, Inc., 2012).

18. Ciresan, D. C., Meier, U., Gambardella, L. M. & Schmidhuber, J. Convolutional Neural Network Committees for Handwritten Character Classification. in 2011 International Conference on Document Analysis and Recognition 1135–1139 (2011).

19. Karpathy, A., Toderici, G., Shetty, S. & Leung, T. Large-scale video classification with convolutional neural networks. Proceedings of the (2014).

20. Simonyan, K. & Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv [cs.CV] (2014).

21. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. Proc. IEEE (2015).

22. Malon, C. D. & Cosatto, E. Classification of mitotic figures with convolutional neural networks and seeded blob features. J. Pathol. Inform. 4, 9 (2013).

23. Ciresan, D., Giusti, A., Gambardella, L. M. & Schmidhuber, J. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. in Advances in Neural Information Processing Systems 25 (eds. Pereira, F., Burges, C. J. C., Bottou, L. & Weinberger, K. Q.) 2843–2851 (Curran Associates, Inc., 2012).

24. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 234–241 (Springer, Cham, 2015).

25. Naito, T. et al. Identification and segmentation of myelinated nerve fibers in a cross-sectional optical microscopic image using a deep learning model. J. Neurosci. Methods 291, 141–149 (2017).

26. Mesbah, R., McCane, B. & Mills, S. Deep convolutional encoder-decoder for myelin and axon segmentation. in 2016 International Conference on Image and Vision Computing New Zealand (IVCNZ) 1–6 (2016).

27. Jelescu, I. O. et al. In vivo quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy. Neuroimage 132, 104–114 (2016).

28. He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proc. IEEE (2015).

29. Simard, P. Y., Steinkraus, D. & Platt, J. C. Best practices for convolutional neural networks applied to visual document analysis. ICDAR (2003).

30. Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv [cs.CV] (2017).

31. Ioffe, S. & Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. in International Conference on Machine Learning 448–456 (2015).

32. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J. Mach. Learn. Res. 15, 1929–1958 (2014).

33. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. arXiv [cs.LG] (2014).

34. Kayalioglu, G. Chapter 10 - Projections from the Spinal Cord to the Brain. in The Spinal Cord 148–167 (Academic Press, 2009).

35. Watson, C. & Harvey, A. R. Chapter 11 - Projections from the Brain to the Spinal Cord. in The Spinal Cord 168–179 (Academic Press, 2009).

36. Schwartz, E. D. et al. Ex vivo evaluation of ADC values within spinal cord white matter tracts. AJNR Am. J. Neuroradiol. 26, 390–397 (2005).

37. Nunes, D., Cruz, T. L., Jespersen, S. N. & Shemesh, N. Mapping axonal density and average diameter using non-monotonic time-dependent gradient-echo MRI. arXiv [physics.med-ph] (2016).

38. Chomiak, T., Hu - PloS one, B. & 2009. What is the optimal value of the g-ratio for myelinated fibers in the rat CNS? A theoretical approach. journals.plos.org (2009).

39. Oquab, M., Bottou, L., Laptev, I., Sivic - Proceedings of the IEEE, J. & 2014. Learning and transferring mid-level image representations using convolutional neural networks. cv-foundation.org (2014).

40. Cohen-Adad, J. et al. (2017, October 25). White Matter Microscopy Database. http://doi.org/10.17605/OSF.IO/YP4QG.

5.7 Acknowledgements

The authors would like to thank Ariane Saliani and Tanguy Duval for helping with the acquisition of SEM data, Dr. Hugues Leblond for providing the human spinal cord sample, Nafisa Husein and Harris Nami for helping with the ground truth labeling of samples, Drs. Adriana Romero Soriano and Yoshua Bengio (MILA - Montreal Institute for Learning Algorithms), and Dr. Robert Brown (McGill - Montreal Neurological Institute) for fruitful discussions on the network design, Drs. Nikola Stikov and Jennifer Campbell for sharing TEM data of macaque, and Dr. Els Fieremans for sharing TEM data of mice. The authors would also like to thank Compute Canada and Calcul Québec for access to computation units and the "NVIDIA Corporation" for offering a Tesla GPU.

This study was funded by the Canada Research Chair in Quantitative Magnetic Resonance Imaging (JCA), the Canadian Institute of Health Research [CIHR FDN-143263], the Canada Foundation for Innovation [32454, 34824], the Fonds de Recherche du Québec - Santé [28826], the Fonds de Recherche du Québec - Nature et Technologies [2015-PR-182754], the Natural Sciences and Engineering Research Council of Canada [435897-2013], IVADO, TransMedTech and the Quebec BioImaging Network.

5.8 Author contributions

AZ, MW and JCA wrote the paper. AZ, MW, VH, PLA and CSP designed and developed the software. JCA supervised the project and provided expert guidance. All authors reviewed the manuscript.

CHAPTER 6 SUPPLEMENTARY METHODS AND RESULTS

The following chapter presents supplementary methods and results that are not included in the two papers of the previous chapters.

6.1 Notes on the *AxonDeepSeg* architecture

The *AxonDeepSeg* architecture presented in Chapter 5 (see **Figure 5.3**) can be also seen as an encoder-decoder architecture (see **Figure 6.1**). This type of representation for fully convolutional networks is also referenced and described in other papers [82][83][59]. The network can be divided into two parts: the convolutional encoder (left, equivalent of the contraction path) and the convolutional decoder (right, equivalent of the expansion path). The task of the encoder is to encode the input image into a more compact and useful representation (i.e. increase number of feature maps but in a lower resolution space). Therefore, the task of the decoder is to perform a semantic projection of the context information learnt by the encoder back into the high-resolution pixel space (i.e. recover the localization information lost during the encoding process).

The fusion between the encoder and the decoder in *AxonDeepSeg* is done by concatenation of the corresponding layers (i.e. merging of the feature maps in the channel axis), but it should be noted that other fusion methods exist. For instance, in the architecture presented by Long and collaborators [41], the combination of the context and localization features is implemented as an addition (i.e. element-wise addition between feature maps).

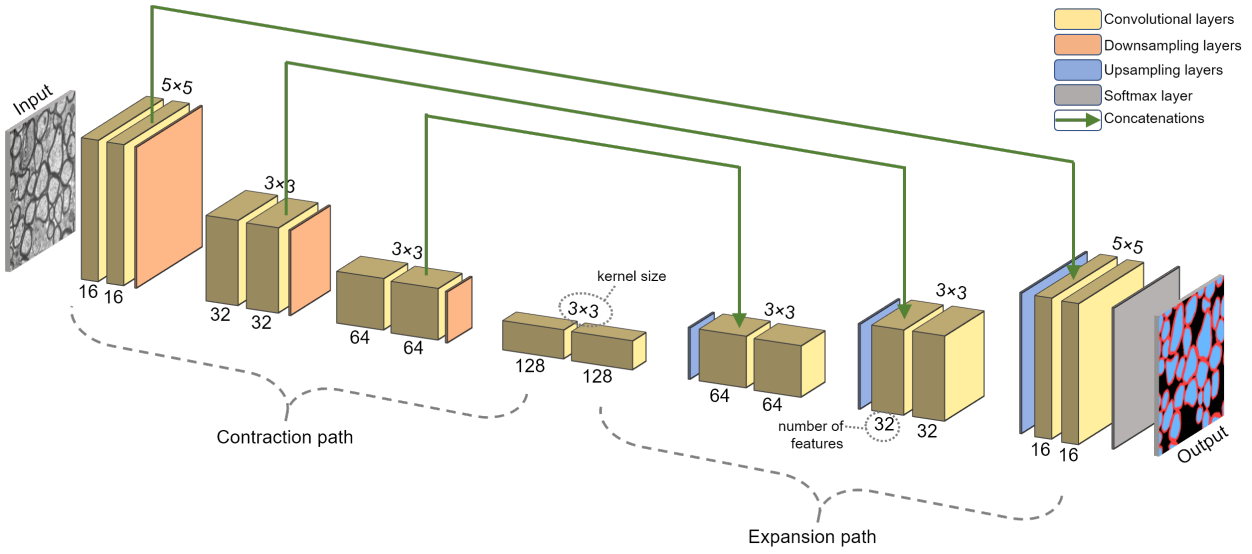

Figure 6.1: The network architecture of *AxonDeepSeg* presented as an encoder-decoder. The contraction path can be seen as a convolutional encoder, while the expansion path can be seen as a convolutional decoder. The architecture of the TEM model is illustrated here.

6.2 Robustness to noise, blurring and intensity changes

One of the main challenges when working with image acquisition systems is to avoid distortions and/or artifacts that can affect the image quality. In microscopy, various types of noise or distortions can occur and cause degradation of the image quality, resulting in poor performance by segmentation algorithms. It has been shown that deep neural networks are sensitive to change in image quality [84], even though methods such as dropout [85] and data augmentation [86] are commonly used during training to improve generalisation and avoid overfitting on training datasets. This is an important consideration given that in our application, microscopy images to be segmented can come from various centres, use different preparation, fixation or acquisition protocols, and/or be acquired by different systems.

6.2.1 Simulation of noise, blurring and intensity changes models

Electron microscopy images can present several distortions (e.g. Gaussian noise [87], blurring [88], illumination inhomogeneities [89]). To assess the robustness of *AxonDeepSeg* segmentation models to image distortions, several distortion models were designed (see **Table 6.1**). For each of the 2 modalities of interest (SEM and TEM), varying levels of distortions were

simulated on a test sample. Degraded samples were then segmented, and the performance was assessed by computing the axon Dice, the myelin Dice and the 3-class pixel-wise accuracy.

Table 6.1: Summary of distortions simulated to assess the robustness of *AxonDeepSeg* models. Additive and multiplicative Gaussian noise, Gaussian blurring, lower contrast, lower brightness and higher brightness were simulated on the test images.

	Type of distortion	**Implementation details**
Noise	Additive Gaussian noise	Implemented as an addition between the input image and an array (of same size) of randomly generated Gaussian values of μ=0 and σ=[0-30].
	Multiplicative Gaussian noise	Implemented as an element-wise multiplication between the input image and an array (of same size) of randomly generated Gaussian values of μ=1 and σ=[0-0.5].
Blurring	Gaussian blurring	Implemented as Gaussian filtering of the input image using σ=[0-10] (and kernel size updated accordingly).
Intensity changes	Lower contrast	Implemented as a weighted linear interpolation (with blending factor *a* varying between 0 and 0.9) between the image and a gray image (all values = average intensity of input image): $\mathrm{Output} = \mathrm{a(Gray)} + (1 - \mathrm{a})\mathrm{Input}$
	Lower brightness	Implemented as a weighted linear interpolation (with blending factor *a* varying between 0 and 0.9) between the image and an empty image (all values = 0): $\mathrm{Output} = \mathrm{a(EmptyImage)} + (1 - \mathrm{a})\mathrm{Input}$
	Higher brightness	Implemented as a weighted linear interpolation (with blending factor *a* varying between 0 and 0.9) between the image and a white image (all values = 255): $\mathrm{Output} = \mathrm{a(WhiteImage)} + (1 - \mathrm{a})\mathrm{Input}$

6.2.2 Results on SEM

To assess the robustness of the SEM model to noise/blurring/intensity changes, a rat spinal cord sample with a field-of-view of 108×77µm (pixel size = 0.07µm) was used (labeled as test sample *rat 2* in Table 5.3). Visual results are presented in **Figures 6.2** and **6.3**, while the effect on the performance metrics is presented in **Figure 6.4**.

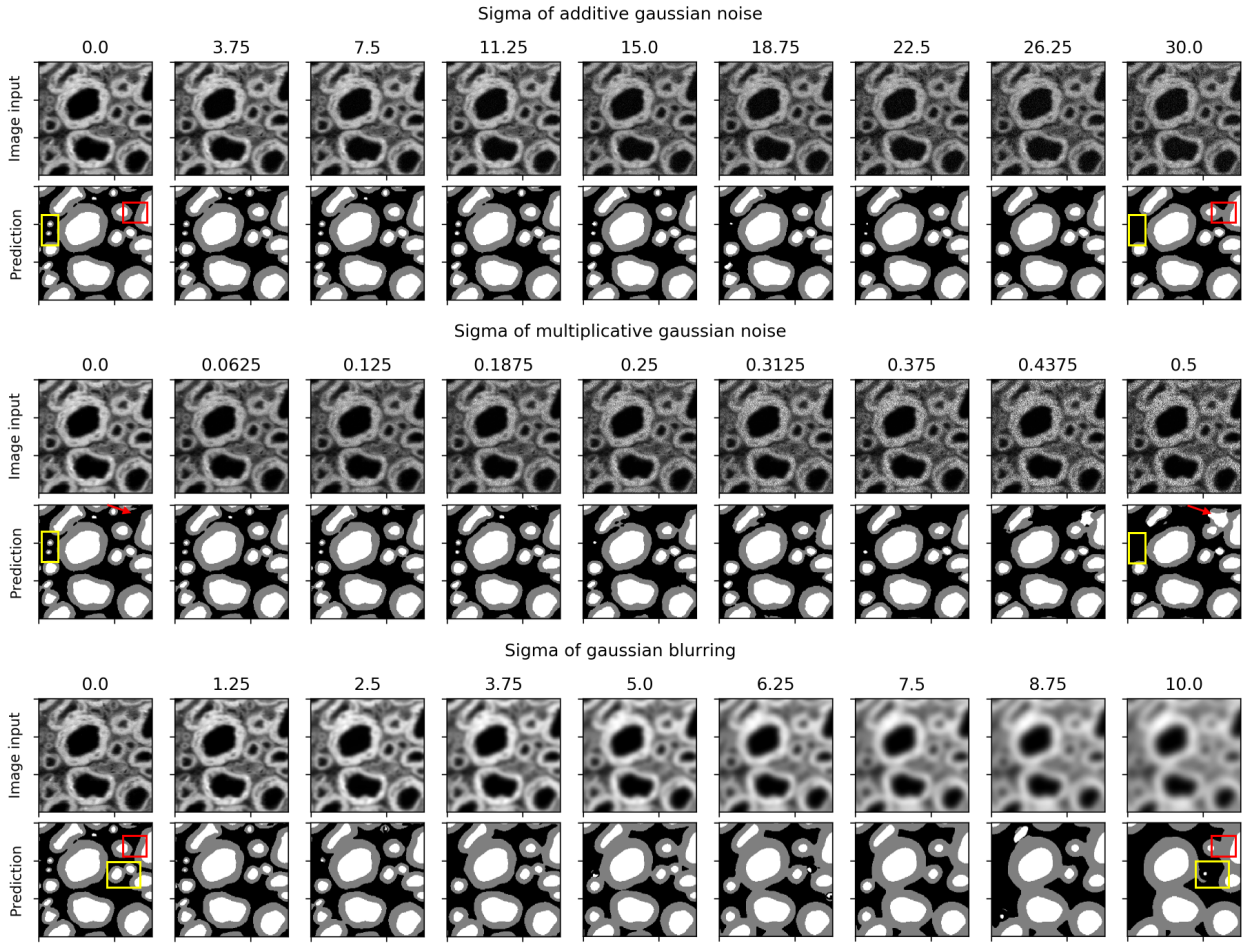

Figure 6.2. Visualization of noise/blurring effects on a small patch of the SEM test sample and on the corresponding prediction maps (myelin=gray, axon=white, background=black) after segmentation. Top: additive gaussian noise of μ=0 and σ varying between 0 and 30. Center: multiplicative gaussian noise of μ=1 and σ varying between 0 and 0.5. Bottom: gaussian blurring of σ between 0 and 10. Yellow squares point out false negatives, red squares point out fusion of neighboring myelin sheaths and red arrows point out false positives.

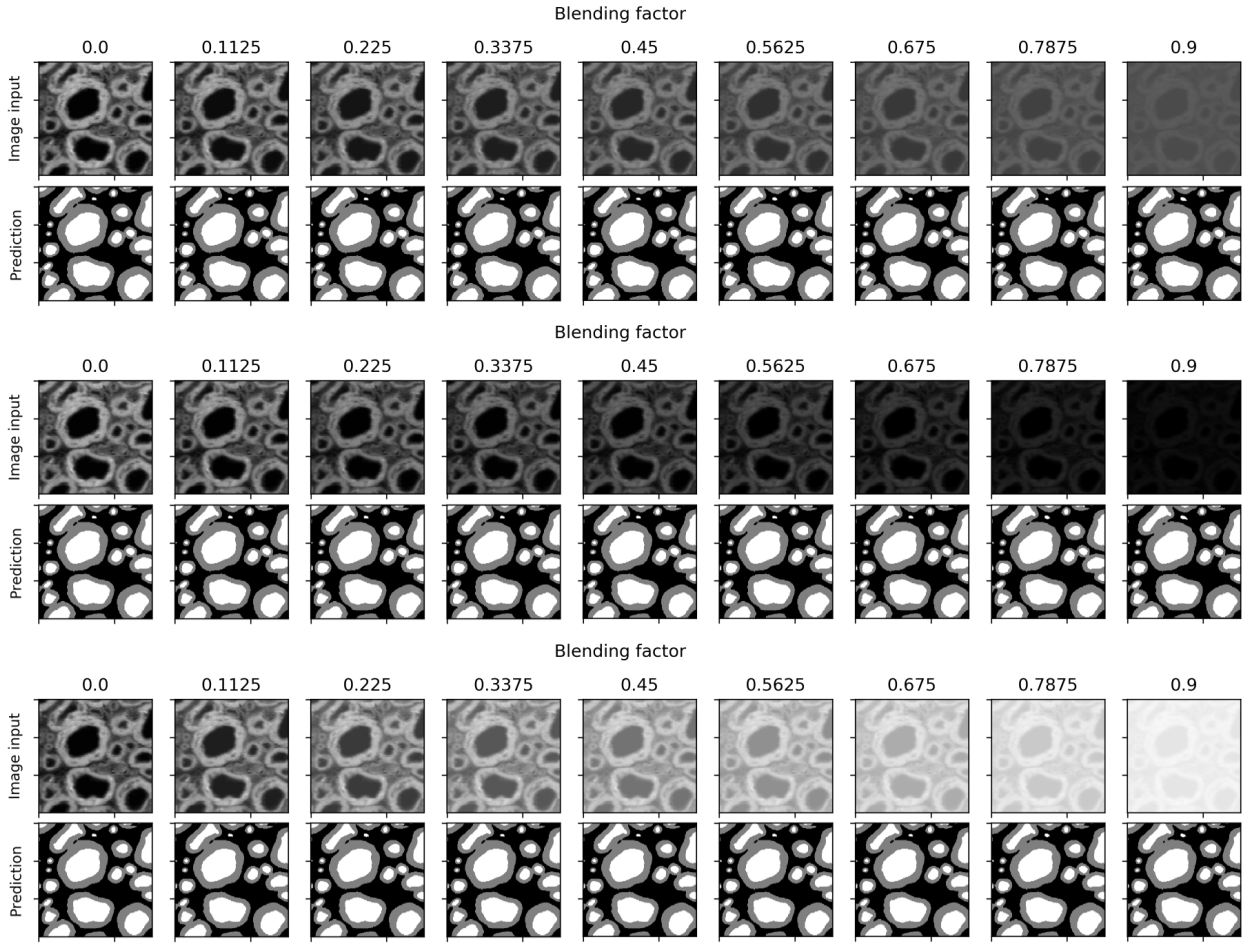

Figure 6.3. Visualization of intensity changes effects on a small patch of the SEM test sample and on the corresponding prediction maps (myelin=gray, axon=white, background=black) after segmentation. Top: image contrast lowered using a blending factor between 0 and 0.9. Center: brightness lowered using a blending factor between 0 and 0.9. Bottom: brightness increased using a blending factor between 0 and 0.9.

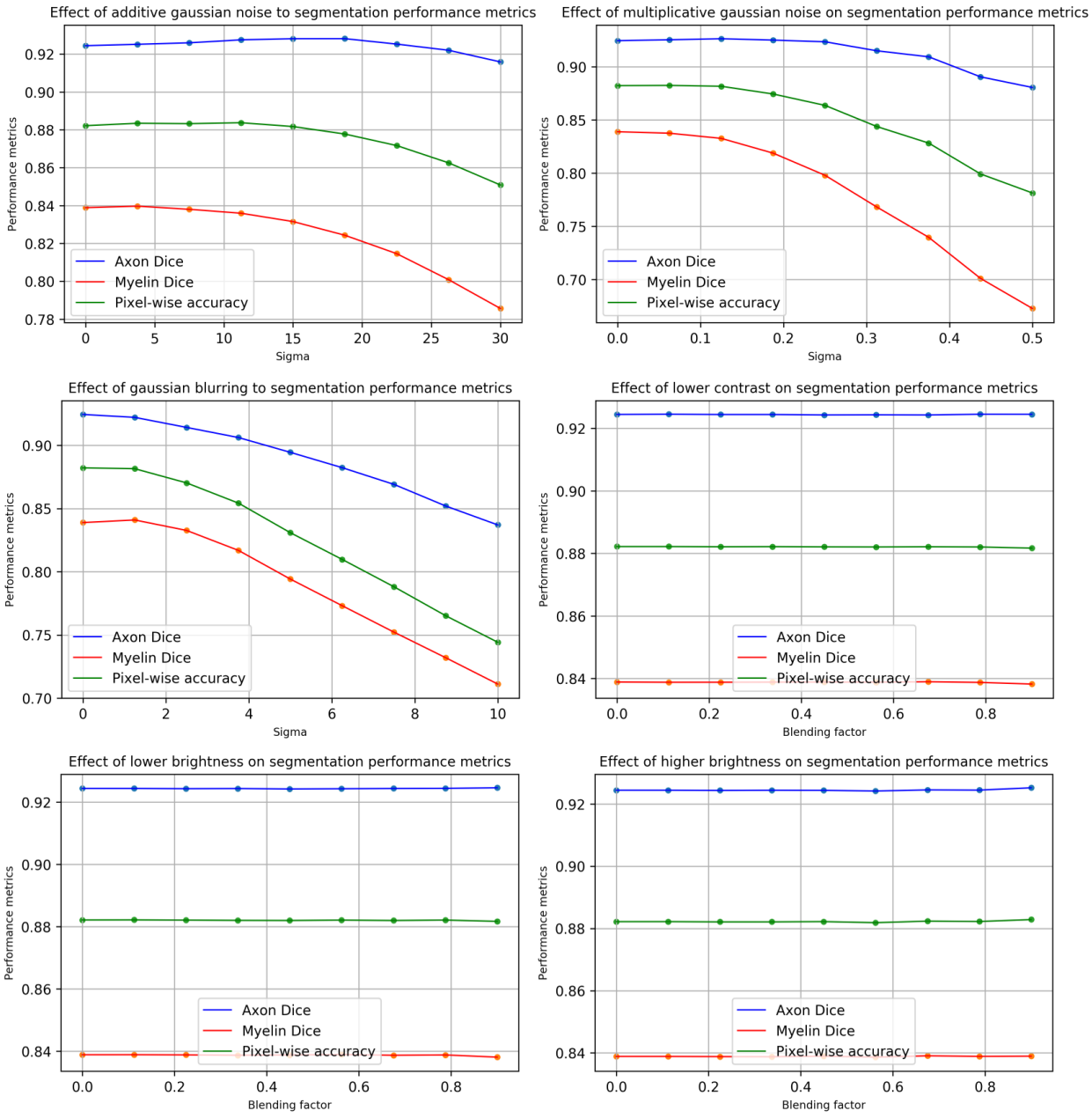

Figure 6.4. Effect of noise/blurring/intensity changes on segmentation of a SEM test sample. Axon Dice, myelin Dice and 3-class pixel-wise accuracy are presented for the following simulations: additive Gaussian noise, multiplicative Gaussian noise, Gaussian blurring, lower contrast, lower brightness, higher brightness.

6.2.3 Results on TEM

To assess the robustness of the TEM model to noise/blurring/intensity changes, a mouse corpus callosum sample with a field-of-view of 6×9µm (pixel size = 0.002µm) was used (one of the 24 test samples of mice in Table 5.3). Visual results are presented in **Figures 6.5** and **6.6**, while the effect on the performance metrics is presented in **Figure 6.7**.

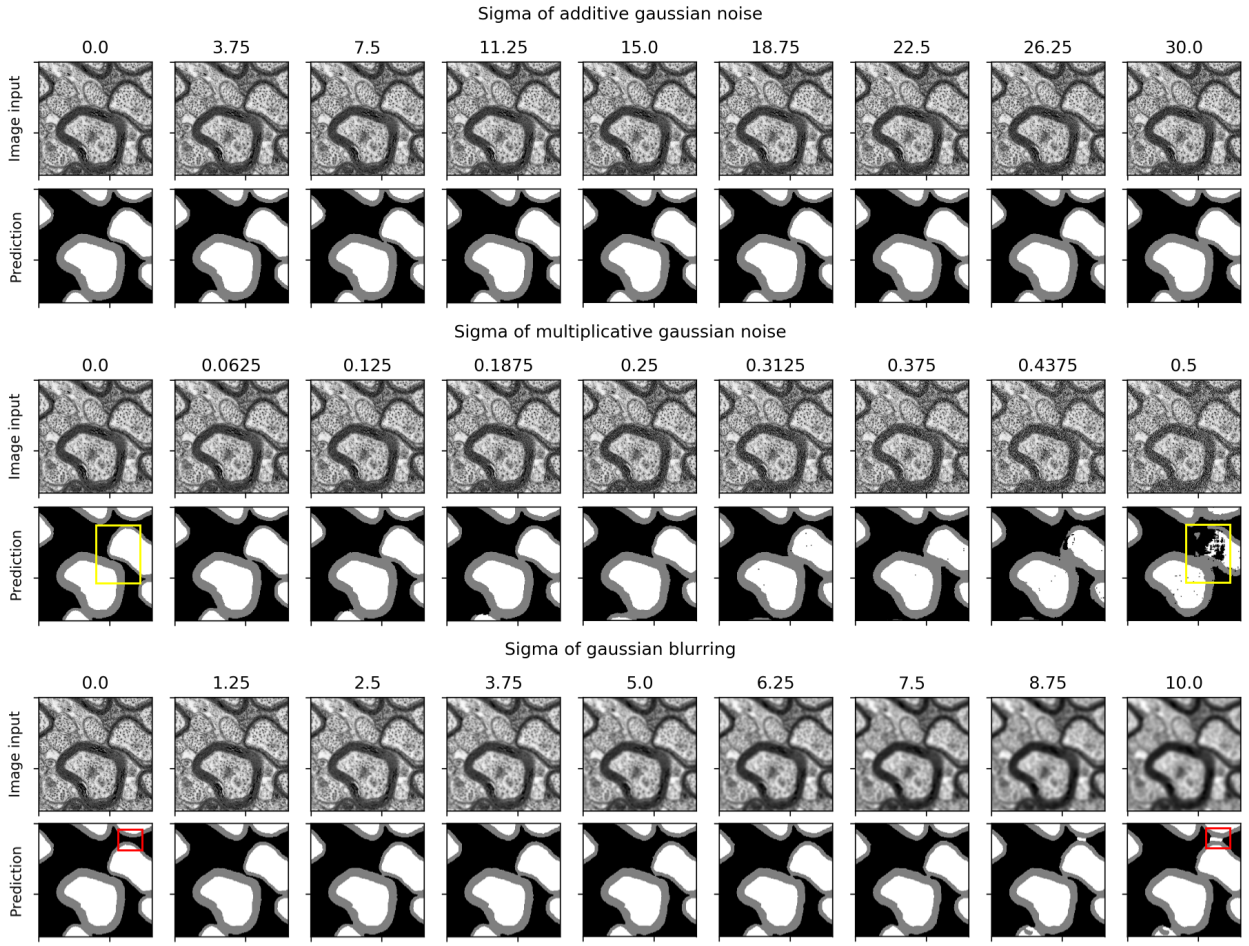

Figure 6.5. Visualization of noise/blurring effects on a small patch of the TEM test sample and on the corresponding prediction maps (myelin=gray, axon=white, background=black) after segmentation. Top: additive gaussian noise of µ=0 and σ varying between 0 and 30. Center: multiplicative gaussian noise of µ=1 and σ varying between 0 and 0.5. Bottom: gaussian blurring of σ between 0 and 10. The yellow square points out false negative pixels and the red square points out misclassification between neighbouring myelin sheaths.

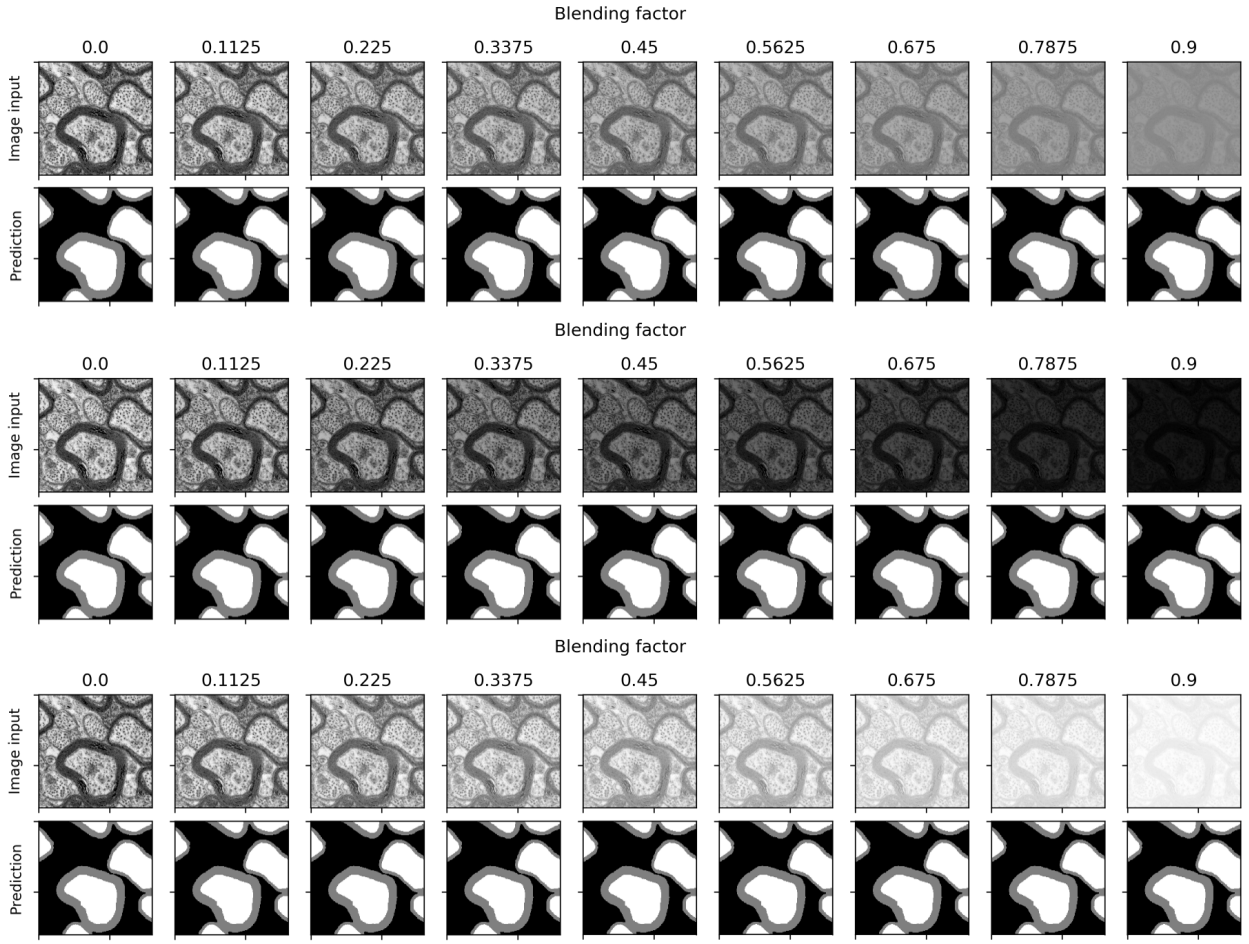

Figure 6.6. Visualization of intensity changes effects on a small patch of the TEM test sample and on the corresponding prediction maps (myelin=gray, axon=white, background=black) after segmentation. Top: lowered contrast using a blending factor between 0 and 0.9. Center: lowered brightness using a blending factor between 0 and 0.9. Bottom: increased brightness using a blending factor between 0 and 0.9.

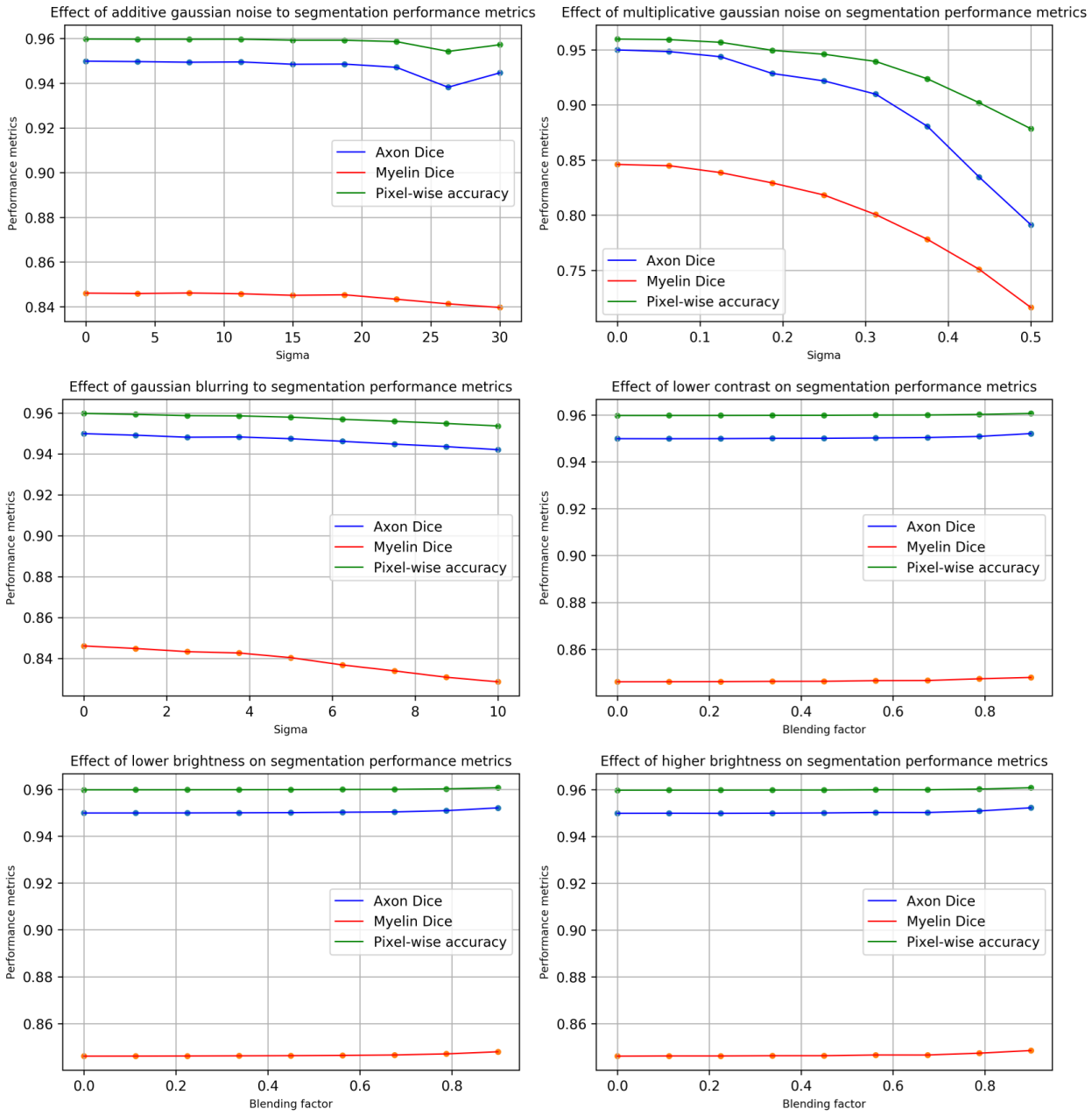

Figure 6.7. Effect of noise/blurring/intensity changes on segmentation of a TEM test sample. Axon Dice, myelin Dice and pixel-wise accuracy are presented for the following simulations: additive Gaussian noise, multiplicative Gaussian noise, Gaussian blurring, lower contrast, lower brightness, higher brightness.

6.2.4 Discussion

Both SEM and TEM models are robust to noise degradation. The SEM model presents a decrease of only 3% of the pixel-wise accuracy when using additive noise of σ=30, and a decrease of 10% when using multiplicative noise of σ=0.5. In the TEM model, decreases of 0.2% and 8% can be observed respectively for additive and multiplicative noise. Visually, the noise degradation can lead to the presence of false positives and false negatives in the segmentation result, as well as errors in the detection of outer boundaries of the myelin sheaths. These results suggest that moderate levels of noise in the input image do not strongly affect the quality of the segmentation.

In blurring experiments, both SEM and TEM models present a decrease in performance (i.e. pixel-wise accuracy of the SEM model drops from 88% to 75% when applying a blurring of σ=10). The presence of blurring in the input can lead to an increase of the number of false negatives for the smaller axons, coarser myelin sheath outer contours and merge between neighbouring myelin sheaths on the segmentation mask (see **Figures 6.2** and **6.5**). It is known that neural networks are very sensitive to blurring [84],[90]. Blurring removes the textures expected in the image (the textures of an image are mostly high frequency information), confusing the network during prediction. It has also been suggested that deeper networks (i.e. more layers) can deal better with noisy inputs (i.e. they can learn more robust features) [84]. However, deeper architectures have more learnable parameters, which could cause overfitting issues if the training dataset is not large enough.

For both SEM and TEM, the trained models are very robust to global intensity changes in the image: the segmentation performance metrics are unaffected by brightness and contrast changes in the input image (see **Figures 6.3** and **6.6**). The robustness to intensity changes is common in CNN architectures [84]. Our architectures also use batch normalization [91], which could further contribute to the robustness to global intensity changes.

6.3 Large set of performance metrics for segmentation

The metrics detailed in the next sub-sections complement the validation metrics already presented in the previous chapters of this thesis. The goal was to develop a more complete set of performance metrics. This can be useful when comparing the performance between trained models.

6.3.1 Pixel-wise performance metrics

A large set of pixel-wise performance metrics was implemented. In the following equations, a true positive (TP) represents a pixel that is true (1) in both the prediction mask and the ground truth mask, while a true negative represents a pixel that is false (0) in both the prediction mask and the ground truth mask. Similarly, a false negative (TN) represents a pixel that is false in the prediction mask but true in the ground truth mask, and a false positive (FP) represents a pixel that is true in the prediction mask and false in the ground truth mask.

The following metrics were implemented for both the axon and myelin masks:

- Sensitivity, computed as TP/(TP+FN);
- Specificity, computed as TN/(TN+FP);
- Precision, computed as TP/(TP+FP);
- False positive rate, computed as FP/(TN+FP), or (1 – Specificity);
- False negative rate, computed as FN/(TP+FN), or (1 – Sensitivity);
- Accuracy, computed as (TP+TN)/(TP+TN+FP+FN);
- Dice, computed as 2TP/(2TP+FP+FN);
- Jaccard, computed as TP/(TP+FP+FN).

The standard Hausdorff distance [92] was also implemented to evaluate the accuracy of the segmentation of the axon-myelin boundary. The interface between the axon and the myelin sheath of each fiber is an important factor in the segmentation (an overestimation of the axon boundary can result in the underestimation of the corresponding myelin sheath inner boundary and add bias to the morphometrics). The implementation consists of first performing a

morphological erosion (using a disk structuring element of size 3) on the prediction and ground truth axonal masks. Then, the overall Hausdorff distance is computed between the binary images of the eroded rings (i.e. the contours removed by the erosion operation).

6.3.2 Axon-wise metrics

The axon-wise Dice is computed as explained in section 4.3.3.4. In brief, for each true positive axon object detected, an individual Dice is computed. Then, a Dice distribution that can be plotted using a boxplot for instance. Furthermore, the nerve fiber detection can be assessed by using the axon sensitivity and precision, as explained in section 4.3.3.3.

An example of use of most of the metrics presented in this section is shown in **Figure 6.8**. Pixel-wise metrics and the axon-wise Dice distribution are presented for a SEM rat spinal cord test sample (the same as the one used for the robustness experiments) segmented using *AxonDeepSeg*.

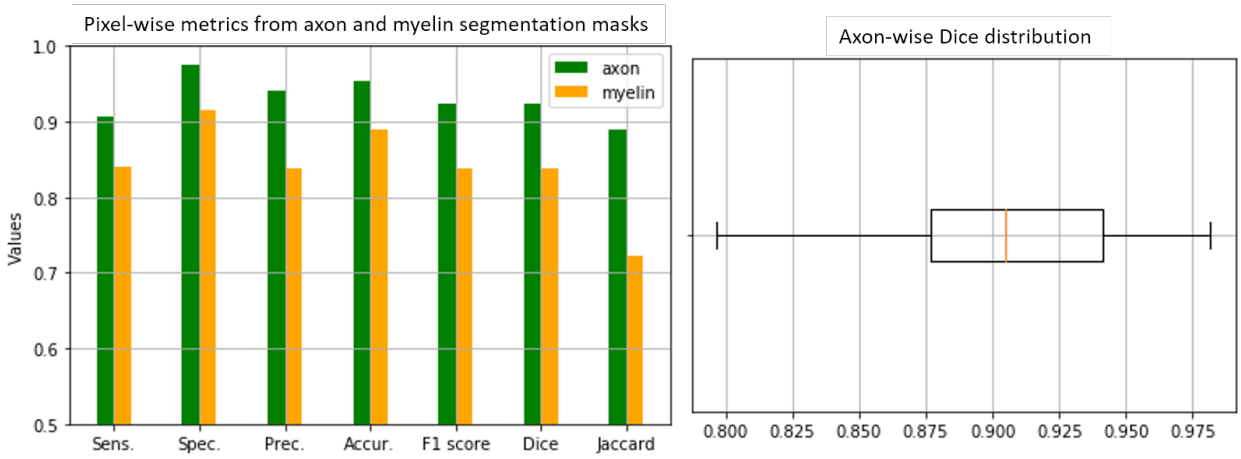

Figure 6.8: Example of pixel-wise (left) and axon-wise (right) segmentation metrics obtained on a SEM test sample. The 25th, 50th and 75th percentiles of the boxplot are 0.877, 0.905 and 0.941 respectively.

6.4 Morphometrics extraction from *AxonDeepSeg*

The following results complement the morphometrics extraction section of the *AxonDeepSeg* paper. **Figure 6.9** shows an example of extraction of individual morphometrics from two different regions of the full rat spinal cord segmentation presented in section 5.4.2.

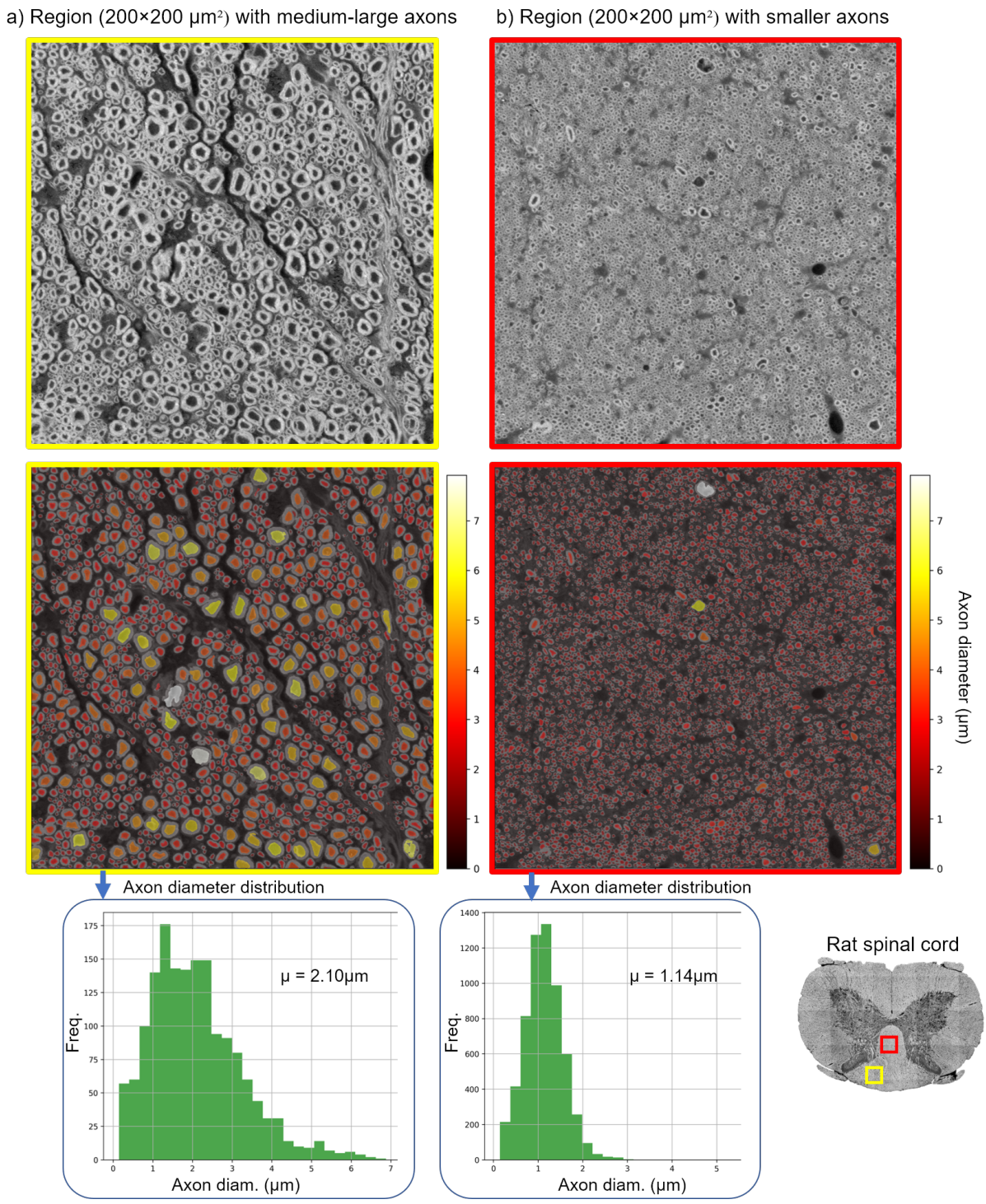

Figure 6.9: Extraction of axon diameter from a region containing medium-large axons (a) and a region containing small axons (b) from a rat spinal cord slice after segmentation with *AxonDeepSeg*. The segmentation results are color-coded for axon diameter. Histogram distributions of axon diameters from each region are presented.

Two regions of 200×200μm² are considered: a region containing axons of medium-large size (extracted from the gracile fasciculus tract of the rat spinal cord) and a region containing axons of smaller size (extracted from the corticospinal tract). Overlays of the segmentation, color-coded for axon equivalent diameter, are presented for both regions. The axon diameter distribution is obtained from the segmented regions. As expected, the gracile fasciculus region presents a larger range of diameters (mean of 2.10 μm), while the corticospinal region contains mostly axons between 0.2 and 2 μm (mean of 1.14 μm). These results suggest that *AxonDeepSeg* is sensitive to all nerve fiber sizes.

6.5 Effect of patch overlap value

One of the challenges of CNN architectures is the management of the segmentation at the borders of the patch. As explained in section 5.3.7, the *AxonDeepSeg* pipeline uses an overlap between patches to give more context information to the network at the borders.

The effect of the overlap value on the segmentation was assessed by experimenting with various overlap values. Prediction was performed on a SEM test sample (the same as the one used for the robustness experiments) by using an overlap value between 1 and 100 pixels. Note that the overlap range tested was chosen to ensure the same number of patches so that the time computation stays the same (i.e. the test image is always divided into 6 patches for all overlap values between 1 and 100). The 3-class pixel-wise accuracy, the axon Dice and the myelin Dice were used to monitor the performance.

6.5.1 Results

Results of the effect of the patch overlap value on the segmentation output of a SEM test sample are presented in **Figure 6.10**. It can be noted that a larger overlap used at prediction time can slightly improve the segmentation results. Increasing the overlap from 1 to 100 pixels increases the axon Dice, myelin Dice and 3-class pixel-wise accuracy by ~0.15%, ~0.12%, and ~0.15% respectively. For this particular test sample, one way of making use of that slight increase in performance while keeping approximately the same computation time (i.e. same number of patches) would be to determine the highest overlap value that gives 6 patches at prediction time.

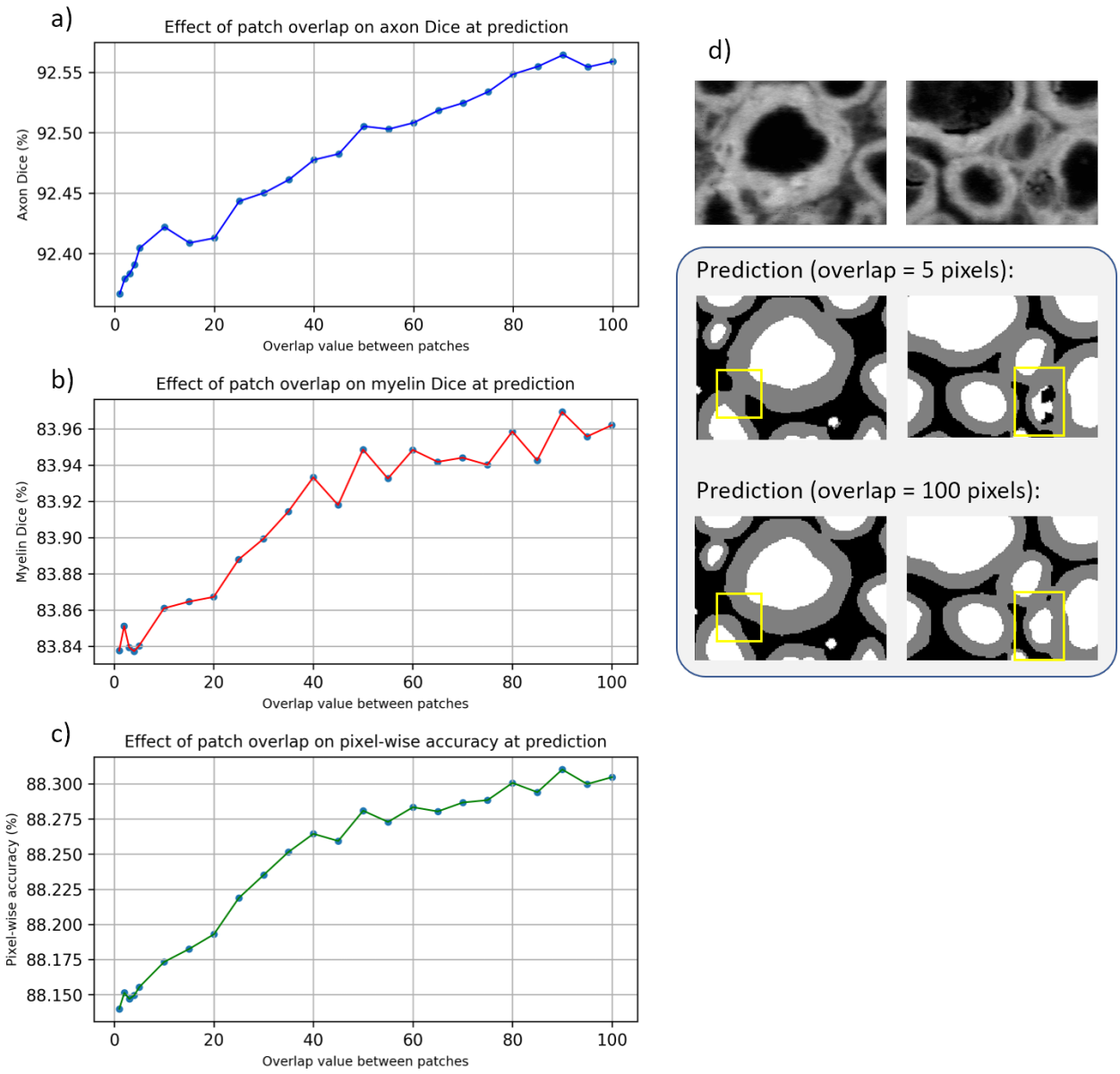

Figure 6.10: Effect of the patch overlap value (in pixels) on the segmentation output of a SEM test sample. For each overlap value, the axon Dice (a), the myelin Dice (b) and the 3-class pixel-wise accuracy (c) are reported. Some examples of visual differences (d) in the segmentation output are displayed (axon=white, myelin=gray, background=black).

6.6 Application to CARS microscopy

Considering that the contrast/features of CARS microscopy samples are similar to those of the SEM samples (i.e. axons are dark and myelin sheaths are white), the idea was to test the SEM model on a CARS test sample. **Figure 6.11** presents the segmentation results on a rat spinal cord test sample compared against manual labeling.

All three main performance metrics were higher than 80%: the axon Dice was 96.2%, the myelin Dice was 83.8% and the 3-class pixel-wise accuracy was 87.6%. Visually, the segmentation and manually labeled masks are very similar. These results are very encouraging and suggest that training a multimodal SEM-CARS model may be an option in future *AxonDeepSeg* model trainings. These results also demonstrate the robustness of the SEM model and its ability to generalize well.

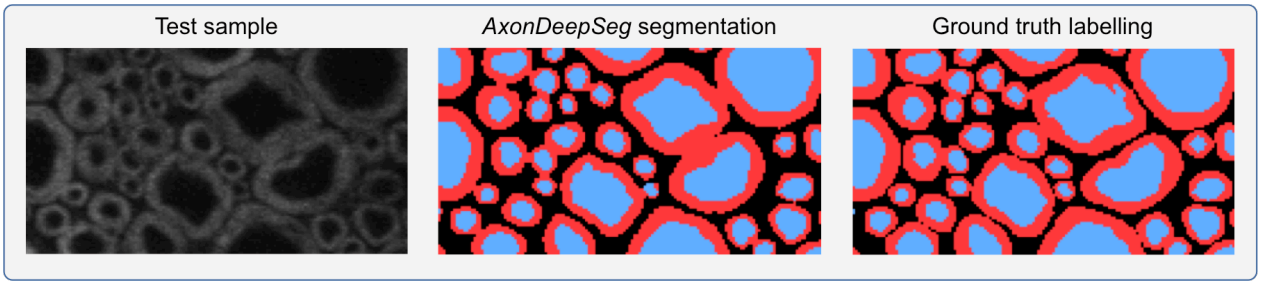

Figure 6.11: Segmentation of a CARS test sample (rat spinal cord) using the SEM model of *AxonDeepSeg*. The test sample, the *AxonDeepSeg* result, and the ground truth labeling are presented.

6.7 Comparison between *AxonSeg* and *AxonDeepSeg*

This project led to the development of two main segmentation frameworks: *AxonSeg* and *AxonDeepSeg*. *AxonSeg* is based on standard image processing algorithms, while *AxonDeepSeg* uses convolutional neural networks. One important goal of this thesis is to compare the segmentation results between the two frameworks.

6.7.1 Method

One rat spinal cord test sample (SEM) and one mouse corpus callosum test sample (TEM) were used (the same test samples as those used for the robustness experiments). Both test images were first segmented with the standard *AxonSeg* pipeline (see section 4.4), and then with the corresponding *AxonDeepSeg* model (either SEM or TEM). Pixel-wise validation metrics of both axon and myelin masks were used to assess the quality of the segmentation.

To evaluate the performance of both segmentation frameworks regarding the morphometrics extraction, the axon volume fraction (AVF), the myelin volume fraction (MVF) and the aggregate g-ratio were extracted from the segmentation results for both SEM and TEM test samples. The morphometrics obtained from *AxonSeg* and *AxonDeepSeg* segmentation maps were compared to those obtained from the ground truth masks.

6.7.2 Results

Figure 6.12 presents pixel-wise axon and myelin segmentation metrics obtained in the SEM test sample and **Figure 6.13** presents the same metrics for the TEM test sample. Distributions of the axon-wise Dice values obtained in the SEM and TEM test samples are presented in **Figure 6.14**. **Figure 6.15** presents the visual results of the segmentations on a small region of the test samples. Results of morphometrics extraction from *AxonSeg*, *AxonDeepSeg* and the ground truth masks from SEM and TEM test samples are presented in **Table 6.2**.

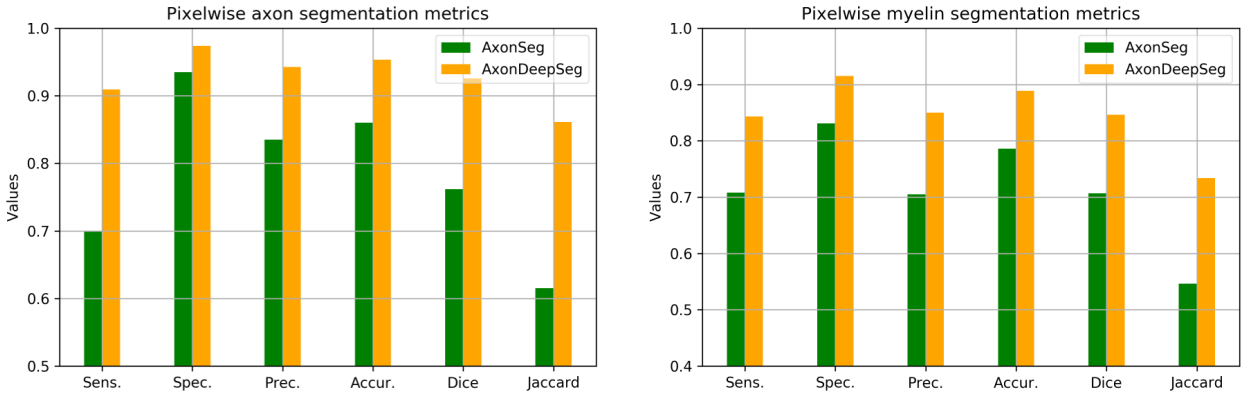

Figure 6.12: Comparison of pixel-wise segmentation performance metrics between *AxonSeg* and *AxonDeepSeg* on a SEM test sample of rat spinal cord. Axon metrics (left) and myelin metrics (right) are presented.

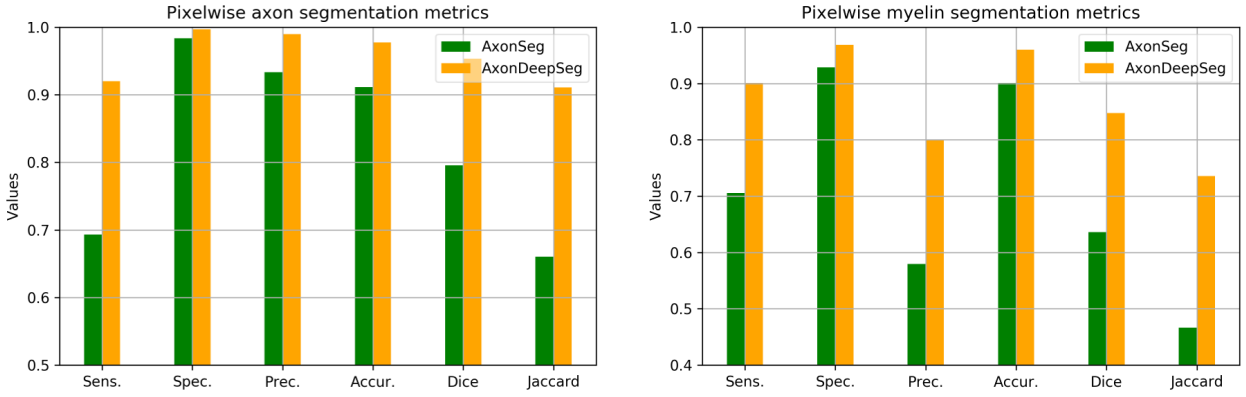

Figure 6.13: Comparison of pixel-wise segmentation performance metrics between *AxonSeg* and *AxonDeepSeg* on a TEM test sample of mouse corpus callosum. Axon metrics (left) and myelin metrics (right) are presented.

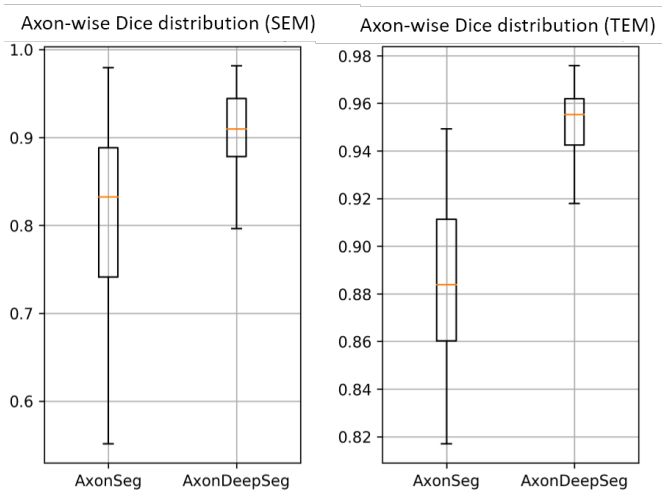

Figure 6.14: Comparison of axon-wise Dice distributions between *AxonSeg* and *AxonDeepSeg*. Left: SEM rat spinal cord test sample. Right: TEM mouse corpus callosum test sample.

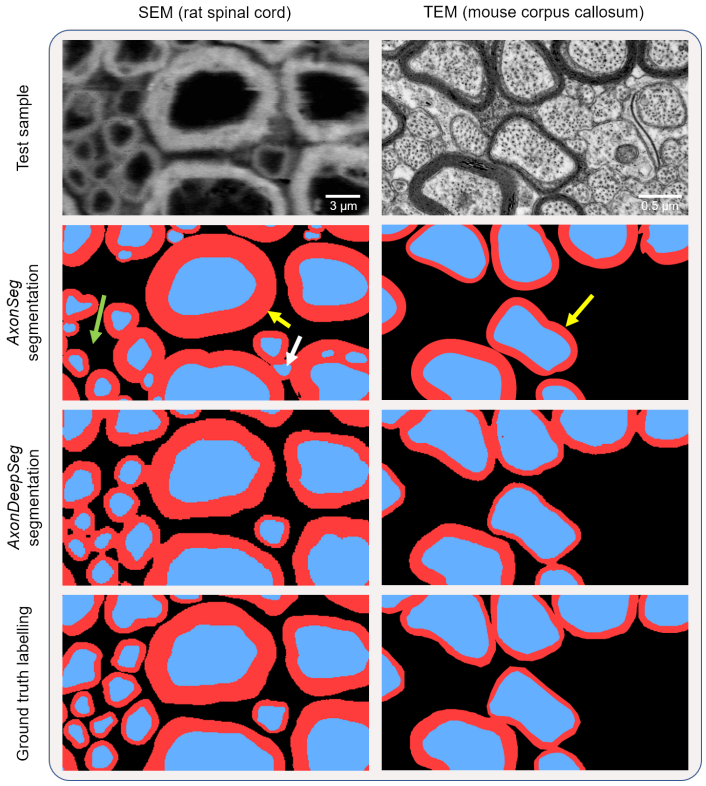

Figure 6.15: Visual comparison between the *AxonSeg* and *AxonDeepSeg* segmentations of a SEM (rat spinal cord) test sample and a TEM (mouse corpus callosum) test sample. Only a small region of the test samples is showed for clarity. Yellow arrows indicate overestimated myelin sheath, green arrows indicate the presence of false negatives, and white arrows indicate the presence of false positives.

Table 6.2: Comparison of aggregate morphometrics extraction from SEM and TEM test samples obtained from *AxonSeg* and *AxonDeepSeg* segmentations and ground truth labelings. For each segmentation method, the axon volume fraction (AVF) and the myelin volume fraction (MVF) of the test samples are computed and reported. The aggregate g-ratio is estimated from the AVF and MVF values, based on [93].

Modality	**Mask used for the morphometrics extraction**	**AVF**	**MVF**	**G-ratio**
SEM	*AxonSeg*	0.2681	0.3650	0.6507
	AxonDeepSeg	0.3083	0.3602	0.6791
	Ground truth labeling	**0.3196**	**0.3632**	**0.6842**
TEM	*AxonSeg*	0.1849	0.1493	0.7438
	AxonDeepSeg	0.2315	0.1380	0.7915
	Ground truth labeling	**0.2490**	**0.1226**	**0.8186**

Overall, it can be observed that the *AxonDeepSeg* segmentations present better performance metrics than *AxonSeg*. In all pixel-wise metrics, *AxonDeepSeg* metrics are higher for both axon and myelin segmentation. In some of the metrics, *AxonDeepSeg* values are between 10-20% higher than the corresponding *AxonSeg* metrics. Similarly, axon-wise Dice distributions obtained from *AxonDeepSeg* present higher medians (~10% higher) and a lower dispersion when compared to *AxonSeg*. Visually, segmentation masks from *AxonDeepSeg* present a lower number of false positives and false negatives and more precise segmentation contours.

Regarding the morphometrics extraction, Table 6.2 suggests that the morphometrics obtained from the *AxonDeepSeg* segmentations are more accurate (i.e. values closer to the ones obtained from the ground truth segmentations). In all cases, it can be observed that the *AxonSeg* segmentation underestimates the axon volume fraction (i.e. more false negative axon pixels in the *AxonSeg* segmentation). This result is in agreement with the large discrepancy between the pixel-wise axon sensitivity values of *AxonSeg* and *AxonDeepSeg* observed in Figures 6.12 and 6.13 (more than 20% difference between the *AxonSeg* and *AxonDeepSeg* axon pixel-wise sensitivities). As illustrated in Figure 6.15, the *AxonSeg* segmentation contains more false

negatives (see green arrow). Moreover, it can also be observed that the *AxonSeg* segmentation overestimates the myelin volume fraction (i.e. more false positive myelin pixels in the *AxonSeg* segmentation). This is also illustrated by the large discrepancy between the pixel-wise myelin precision values of *AxonSeg* and *AxonDeepSeg* observed in Figures 6.12 and 6.13, especially in the TEM test sample.

The underestimation of the axon volume fraction and the overestimation of the myelin volume fraction in the *AxonSeg* segmentations result in an underestimated aggregate g-ratio. For instance, in the SEM test sample, the *AxonSeg* g-ratio is almost 5% lower than the g-ratio obtained from the ground truth labeling, while the *AxonDeepSeg* g-ratio is only 0.75% lower than the one computed from the ground truth segmentation. Overall, these results suggest that the *AxonDeepSeg* segmentations lead to more accurate morphometrics computations.

In terms of computational time, the *AxonDeepSeg* prediction pipeline is significantly faster. For instance, the same SEM test sample took ~100 seconds to segment with the *AxonSeg* pipeline, while the prediction time was ~27 seconds with *AxonDeepSeg* (i.e. ~4× faster). Further discussion between the two frameworks is presented in Chapter 7.

CHAPTER 7 GENERAL DISCUSSION

This general discussion complements the discussions presented in the two papers (see sections 4.5 and 5.5) and summarizes the project regarding the achievement of the specific objectives identified in section 2.7. The main contributions of this project were the design and implementation of frameworks (*AxonSeg* and *AxonDeepSeg*) capable of performing robust axon and myelin segmentation and extraction of morphological metrics.

7.1 Achievement of the objectives

Chapter 4 presented the first paper, titled "*AxonSeg: open source software for axon and myelin segmentation and morphometric analysis*". The first specific objective was achieved by proposing a segmentation pipeline that combines preliminary axon segmentation (extended-minima), axon discrimination based on shape and intensity features (discriminant analysis), preliminary myelin segmentation (edge detection) and segmentation refinement (active contours). To achieve the second specific objective, the pipeline was validated on SEM, CARS and OM samples of rat and cat samples. The last specific objective was achieved by implementing an extraction pipeline and applying it to document morphology of white matter tracts of a cat spinal cord.

Chapter 5 presented the second paper, titled "*AxonDeepSeg: automatic axon and myelin segmentation from microscopy data using convolutional neural networks*". The first specific objective was achieved by implementing a convolutional neural network for the 3-class semantic segmentation of axon and myelin. The second specific objective was achieved by training SEM and TEM models and validating on various tissues (spinal cord and corpus callosum) and various species (rat, human, mouse, macaque). All 3-class pixel-wise accuracies on test samples were higher than 81%. The last specific objective was achieved by implementing a morphometrics extraction pipeline from the prediction output. The morphometrics extraction pipeline was validated by comparing morphometrics obtained from rat spinal cord tracts with the literature.

Although all three specific objectives of this research project were achieved by both software packages (*AxonSeg* and *AxonDeepSeg*), *AxonDeepSeg* achieved better results regarding the objectives, as shown in section 6.7.

7.2 Comparison between *AxonSeg* and *AxonDeepSeg*

At the beginning of the project, a 2-class (axon and non-axon) deep learning module was meant to be used as a replacement of the axon segmentation step of *AxonSeg* (i.e. axon segmented by a convolutional neural network and then myelin segmented by the *AxonSeg* pipeline). However, it turned out that the end-to-end 3-class approach currently implemented in *AxonDeepSeg* was a better alternative for several reasons.

First, having an end-to-end pipeline offers a better automation and is more convenient to use. As literature review pointed out, a combination of standard image processing segmentation algorithms is often necessary because the axon-myelin segmentation is a complex task. This means that most standard segmentation frameworks will have complex pipelines consisting of several steps. For instance, in *AxonSeg*, the segmentation pipeline consists of the four main steps: preliminary axon segmentation, axon discrimination, preliminary myelin sheath identification, and refinement of boundaries with active contours. This type of pipeline can also lead to an accumulation of segmentation errors, as an erroneous output of a step will be the input of the following step (e.g. a false positive axon will lead to the segmentation of an inexistent myelin sheath).

Secondly, a deep learning framework automates the feature engineering part, since the network itself learns the relevant features for a given application. In *AxonSeg*, the shape and intensity features of the discriminant analysis module were selected among a large set of possible features after extensive exploration. In *AxonDeepSeg*, this is implicitly done by the learned filters of the convolutional layers.

Finally, *AxonDeepSeg* allows better control of the computation time at prediction. The prediction time in *AxonDeepSeg* for a given model is only proportional to the size of the sample (i.e. number of patches to segment), while the segmentation time in *AxonSeg* depends on both the sample size and the number of axons to segment in the sample (i.e. for each axon object detected, the myelin sheath is coarsely identified, then refined using active contours).

The points mentioned above and the superior segmentation results obtained by *AxonDeepSeg* indicate that *AxonDeepSeg* will be the main focus of the lab in the long term. Nevertheless, it should be noted that efforts have been put into maintaining compatibility

between *AxonSeg* and *AxonDeepSeg*. For instance, if the myelin segmentation of *AxonDeepSeg* fails, the segmented axon mask can be used as input in *AxonSeg* in the myelin segmentation step of the pipeline. Similarly, most *AxonSeg* tools developed for tasks such as segmentation validation or morphometrics extraction can be used to process the *AxonDeepSeg* output predictions, and vice-versa.

7.3 Training of *AxonDeepSeg* models

The patch size and the pixel size of a trained model are very important parameters and should be carefully designed when training new models. The patch size can be seen as a hyperparameter since it controls the visual context given to the network (for this particular application of multi-instance segmentation, the network has to be given information about the relative positions of axons and myelin and the presence of neighbouring nerve fibers). To ensure that the context information of each patch stays similar between test samples, a target pixel size needs to be chosen. Issues can potentially arise when segmenting samples whose native pixel size is very different from the target pixel size of the model (e.g. factor of 10). Important subsampling can lead to loss of information but also drastically reduce the size of the image, possibly causing size issues when fed into the model (i.e. subsampled sample being smaller than the patch size of the network). Similarly, important upsampling to meet the target pixel size of the model can significantly reduce the context available in a single patch (e.g. only one axon-myelin pair in the patch).

Another important discussion point is the selection of the training dataset. The optimal solution would be to train a single model that performs well on all datasets. However, as experienced when training *AxonDeepSeg* models, training on multimodal datasets (e.g. combination of SEM and TEM) may lead to lower performance on both datasets. At the same time, models trained on a specific dataset (e.g. SEM images acquired with the same protocol by the same research group) will usually perform well but may present overfitting issues. Similarly, the lack of labeled data may also lead to overfitting issues or weak generalization. Novel architectures such as Generative Adversarial Networks (GANs) that can notably be used to generate realistic synthetic data [94]–[97] could be incorporated into the current architecture to improve performance.

A potential issue that can appear in the segmentation result is the merging between neighboring myelin sheaths. Having separate and axon-myelin pairs is desirable and would facilitate the morphometrics extraction pipeline. An elegant way of dealing with this issue is to add weights on the outer borders of the myelin sheaths during the training (i.e. penalize more when the borders are not well segmented). Another way to address this issue could be to use post-processing tools to update the segmentation masks (i.e. first apply a distance transform and then apply a watershed algorithm to separate the touching myelin sheaths).

7.4 Validation methodology

One of the main challenges that arose during the validation methodology was the manual (ground truth) labeling. It is an essential part of the segmentation validation procedure but can often be very time consuming and tedious for this specific application (i.e. manual labeling of multiple axon and myelin contours in a single test sample). Several procedures and tools were explored (e.g. *GIMP*, *Fiji*, *MIPAV*) to identify the most convenient ones and establish a standard procedure in order to avoid variability between labeling done by different people. With *AxonDeepSeg* being based on supervised learning, it means that the performance of the segmentation models strongly depends on the accuracy, precision and number of labeled samples.

7.5 Software distribution and collaboration

A considerable amount of work not featured in this thesis has been put into the development and maintenance of both *AxonSeg* and *AxonDeepSeg* software packages. The long-term sustainability and success of *AxonDeepSeg* relies on the collaboration between the development team and potential users. Shared datasets from collaborators can improve trained models (or build new models), which will then be able to achieve better segmentation results and be helpful to a large number of research groups.

As mentioned previously, both packages have been shared online on *GitHub*. The implementation of open source software requires significant work and introduces new challenges. The *AxonDeepSeg* package can be installed as a command line application, and detailed documentation is provided, for both potential users and future developers. Moreover, a set of *Jupyter* notebooks (i.e. *IPython*) were implemented to facilitate tasks such as segmenting samples, analysing and visualizing performance metrics between different samples or models,

computing and visualizing morphometrics from samples of interest and training new models. The same can be said about the *AxonSeg* package (i.e. documentation and tutorial scripts available for the user). It should be noted that the *AxonDeepSeg* package is more adapted to the open source philosophy, as it is coded in the open source *Python* language and based on the open source *TensorFlow* framework.

CHAPTER 8 CONCLUSION AND RECOMMENDATIONS

The main objective of this research project was to develop an automated framework for the segmentation of axons and myelin sheaths from microscopy data. Two segmentation frameworks were presented: *AxonSeg* and *AxonDeepSeg*. *AxonSeg* is based on standard image processing algorithms and combines the extended-minima algorithm and a discriminant analysis based on shape and intensity features for the axon segmentation, a preliminary myelin edge detection, and a refinement with active contours. *AxonDeepSeg* uses a convolutional neural network that performs an end-to-end 3-class semantic segmentation. Both frameworks were validated and were used to extract morphometric measurements from white matter tracts of the spinal cord. Overall, *AxonDeepSeg* achieved superior segmentation results and offers more advantages based on its design and pipeline.

The software packages developed will likely contribute to the advancement of the neuroscience field. Three major applications can benefit from *AxonSeg* and *AxonDeepSeg*: documentation of nerve morphometry across species and tissues, study of demyelination models, and validation of new MRI biomarkers. Both packages are open source, which is in accordance with the philosophy adopted by NeuroPoly to provide free and open access software to the research community. The feedback from users and collaboration with other research groups should most likely improve the current methods.

Future developments and improvements of the *AxonDeepSeg* framework are possible. Training of new models on other microscopy modalities is planned, as well as training on larger datasets to improve robustness and generalization. Novel CNN architectures can also be experimented. For instance, integration of *Generative Adversarial Networks* into the current architecture could lead to better performance. Another way to possibly improve performance could be to add post-processing tools to the current pipeline (e.g. use morphological operators to close axon/myelin holes or eliminate small debris of false positive pixels). Another interesting development would be to adapt the current 2D architecture to segment 3D microscopy samples. Finally, a short-term priority would be to test the trained models on demyelinated samples to assess the sensitivity to small changes of the myelin thickness.

BIBLIOGRAPHY

[1] "What is MS? — MS Society of Canada." [Online]. Available: https://mssociety.ca/about-ms/what-is-ms. [Accessed: 20-Mar-2018].

[2] "26.1: Anatomy of the Nervous System - Biology LibreTexts." [Online]. Available: https://files.mtstatic.com/site_4463/9181/0?Expires=1524774966&Signature=iqDkUwG~G1HX9UYdfQCY7rxiZOaN2DUhhiM8may-ok7JuC8oXZApG1NpQUsdDGddhKZAMlrX3veLMUFJgplGkZ45PkK98OgB2mIojCKAJJtybwN7e1osUn0CoGwfVrba7B95yawfHWMmy4VHRcNzBQTyJdcoc~NBprmT-U8H4eg_&Key-Pair-. [Accessed: 10-Mar-2018].

[3] G. J. Siegel, *Basic neurochemistry : molecular, cellular, and medical aspects*. Lippincott Williams & Wilkins, 1999.

[4] "Saltatory conduction - The School of Biomedical Sciences Wiki." [Online]. Available: https://teaching.ncl.ac.uk/bms/wiki/index.php/Saltatory_conduction. [Accessed: 10-Mar-2018].

[5] "Unmyelinated and Myelinated Axons • GetBodySmart." [Online]. Available: https://www.getbodysmart.com/nervous-system/unmyelinated-myelinated-axons. [Accessed: 10-Mar-2018].

[6] P. Johns, "Neurons and glial cells," in *Clinical Neuroscience*, Elsevier, 2014, pp. 61–69.

[7] "Conduction in a myelinated nerve fiber: saltatory conduction." [Online]. Available: http://antranik.org/wp-content/uploads/2012/04/conduction-in-a-myelinated-nerve-fiber-saltatory-conduction.jpg. [Accessed: 10-Mar-2018].

[8] "Myelinated neuron." [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/thumb/c/c1/Myelinated_neuron.jpg/220px-Myelinated_neuron.jpg. [Accessed: 10-Mar-2018].

[9] R. D. Fields, "Neuroscience. Change in the brain's white matter.," *Science*, vol. 330, no. 6005, pp. 768–9, Nov. 2010.

[10] A. Fitsiori, D. Nguyen, A. Karentzos, J. Delavelle, and M. I. Vargas, "The corpus callosum: white matter or terra incognita.," *Br. J. Radiol.*, vol. 84, no. 997, pp. 5–18, Jan.

2011.

[11] “White vs Gray Matter - Difference.” [Online]. Available: https://difference.guru/difference-between-white-and-gray-matter/. [Accessed: 10-Mar-2018].

[12] “Corpus callosum.” [Online]. Available: http://web.stanford.edu/group/hopes/cgi-bin/hopes_test/wp-content/uploads/2010/06/corpuscallosum.png. [Accessed: 10-Mar-2018].

[13] “Labelled Diagram Of Brain And Spinal Cord Brain And The Cranial Nerves - Anatomy Inner Body.” [Online]. Available: https://anatomyinnerbody.com/labelled-diagram-of-brain-and-spinal-cord/labelled-diagram-of-brain-and-spinal-cord-brain-and-the-cranial-nerves-2/. [Accessed: 11-Mar-2018].

[14] “Spinal Cord Tracts.” [Online]. Available: https://oli.cmu.edu/jcourse/workbook/activity/page?context=434898b880020ca600ff3a49438dbbba. [Accessed: 11-Mar-2018].

[15] “Tracts of the Spinal Cord.” [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/thumb/b/b2/Spinal_cord_tracts_-_English.svg/874px-Spinal_cord_tracts_-_English.svg.png. [Accessed: 11-Mar-2018].

[16] H. Lassmann, “Multiple Sclerosis Pathology.,” *Cold Spring Harb. Perspect. Med.*, vol. 8, no. 3, p. a028936, Mar. 2018.

[17] D. H. Mahad, B. D. Trapp, and H. Lassmann, “Pathological mechanisms in progressive multiple sclerosis,” *Lancet Neurol.*, vol. 14, no. 2, pp. 183–193, Feb. 2015.

[18] F. Papastefanaki and R. Matsas, “From demyelination to remyelination: The road toward therapies for spinal cord injury,” *Glia*, vol. 63, no. 7, pp. 1101–1125, 2015.

[19] R. Geraldes *et al.*, “The current role of MRI in differentiating multiple sclerosis from its imaging mimics,” *Nat. Rev. Neurol.*, Mar. 2018.

[20] Y. Ge, “Multiple Sclerosis: The Role of MR Imaging.”

[21] S. A. Trip, D. H. Miller, and D. H. Miller, “Imaging in Multiple Sclerosis,” *J Neurol Neurosurg Psychiatry*, vol. 76, pp. 11–18, 2005.

[22] “Multiple sclerosis.” [Online]. Available: http://i.imgur.com/RHADjf6.png. [Accessed: 11-Mar-2018].

[23] Y. Bando *et al.*, “Abnormal morphology of myelin and axon pathology in murine models of multiple sclerosis,” *Neurochem. Int.*, vol. 81, pp. 16–27, Feb. 2015.

[24] K. Schmierer, A. McDowell, N. Petrova, D. Carassiti, D. L. Thomas, and M. E. Miquel, “Quantifying multiple sclerosis pathology in post mortem spinal cord using MRI,” *Neuroimage*, Jan. 2018.

[25] A. Saliani, B. Perraud, T. Duval, N. Stikov, S. Rossignol, and J. Cohen-Adad, “Axon and Myelin Morphology in Animal and Human Spinal Cord,” *Front. Neuroanat.*, vol. 11, p. 129, Dec. 2017.

[26] O. Cuisenaire, “Automatic segmentation and measurement of axons in microscopic images,” *Proc. SPIE*, vol. 3661, pp. 920–929, 1999.

[27] E. Romero, O. Cuisenaire, J. F. Denef, J. Delbeke, B. Macq, and C. Veraart, “Automatic morphometry of nerve histological sections,” *J. Neurosci. Methods*, vol. 97, no. 2, pp. 111–122, 2000.

[28] A. Zaimi, T. Duval, A. Gasecka, D. Côté, N. Stikov, and J. Cohen-Adad, “AxonSeg: open source software for axon and myelin segmentation and morphometric analysis,” *Front. Neuroinform.*, vol. 10, 2016.

[29] “Limit of resolution of optical microscope - WikiLectures.” [Online]. Available: https://www.wikilectures.eu/w/Limit_of_resolution_of_optical_microscope. [Accessed: 11-Mar-2018].

[30] J. Kwon *et al.*, “Label-free nanoscale optical metrology on myelinated axons in vivo,” *Nat. Commun.*, vol. 8, no. 1, p. 1832, Dec. 2017.

[31] “Resolution.” [Online]. Available: https://www.cas.miamioh.edu/mbiws/microscopes/resolution.html. [Accessed: 11-Mar-2018].

[32] M. Müller and A. Zumbusch, “Coherent anti-Stokes Raman Scattering Microscopy,” *ChemPhysChem*, vol. 8, no. 15, pp. 2156–2170, Oct. 2007.

[33] S. Bégin *et al.*, “Automated method for the segmentation and morphometry of nerve fibers in large-scale CARS images of spinal cord tissue.,” *Biomed. Opt. Express*, vol. 5, no. 12, pp. 4145–61, 2014.

[34] X. Xu *et al.*, “Using Coherent Anti-Stokes Raman Scattering (CARS) to Image Brain Tissues,” in *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2007, pp. 480–483.

[35] H. Wang, Y. Fu, P. Zickmund, R. Shi, and J.-X. Cheng, “Coherent Anti-Stokes Raman Scattering Imaging of Axonal Myelin in Live Spinal Tissues,” *Biophys. J.*, vol. 89, no. 1, pp. 581–591, Jul. 2005.

[36] Y. Fu, H. Wang, T. B. Huff, R. Shi, and J.-X. Cheng, “Coherent anti-Stokes Raman scattering imaging of myelin degradation reveals a calcium-dependent pathway in lyso-PtdCho-induced demyelination.,” *J. Neurosci. Res.*, vol. 85, no. 13, pp. 2870–81, Oct. 2007.

[37] J. O’Brien, S. Hoque, D. Mulvihill, and K. Sirlantzis, “Automated Cell Segmentation of Fission Yeast Phase Images - Segmenting Cells from Light Microscopy Images,” in *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, 2017, pp. 92–99.

[38] J. Kong *et al.*, “Automated cell segmentation with 3D fluorescence microscopy images,” in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015, pp. 1212–1215.

[39] S. Dimopoulos, C. E. Mayer, F. Rudolf, and J. Stelling, “Accurate cell segmentation in microscopy images using membrane patterns,” *Bioinformatics*, vol. 30, no. 18, pp. 2644–2651, Sep. 2014.

[40] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[41] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 3431–3440, 2015.

[42] B. Weyn, M. Van Remoortere, R. Nuydens, T. Meert, and G. Van De Wouwer, “A multiparametric assay for quantitative nerve regeneration evaluation,” *J. Microsc.*, vol. 219, no. 2, pp. 95–101, Aug. 2005.

[43] F. Urso-Baiarda and A. O. Grobbelaar, “Practical nerve morphometry,” *J. Neurosci. Methods*, vol. 156, no. 1–2, pp. 333–341, Sep. 2006.

[44] H. L. More, J. Chen, E. Gibson, J. M. Donelan, and M. F. Beg, “A semi-automated method for identifying and measuring myelinated nerve fibers in scanning electron microscope images,” *J. Neurosci. Methods*, vol. 201, no. 1, pp. 149–158, 2011.

[45] Y.-C. Liu *et al.*, “Quantitative Measurement of Nerve Cells and Myelin Sheaths from Microscopic Images via Two-Staged Segmentation,” Springer, Berlin, Heidelberg, 2012, pp. 486–494.

[46] Y.-L. Fok, J. C. K. Chan, and R. T. Chin, “Automated analysis of nerve-cell images using active contour models,” *IEEE Trans. Med. Imag.*, vol. 15, no. 3, pp. 353–368, 1996.

[47] M. Gierthmuehlen *et al.*, “Computational Tissue Volume Reconstruction of a Peripheral Nerve Using High-Resolution Light-Microscopy and Reconstruct,” *PLoS One*, vol. 8, no. 6, p. e66191, Jun. 2013.

[48] X. Zhao, Z. Pan, J. Wu, G. Zhou, and Y. Zeng, “Automatic identification and morphometry of optic nerve fibers in electron microscopy images,” *Comput. Med. Imaging Graph.*, vol. 34, no. 3, pp. 179–184, 2010.

[49] “A Simple Way to Understand Machine Learning vs Deep Learning | Zendesk Blog.” [Online]. Available: https://www.zendesk.com/blog/machine-learning-and-deep-learning/. [Accessed: 12-Mar-2018].

[50] F. Chollet, *Deep learning with Python*, 1st ed. Manning Publications, 2017.

[51] “CS231n Convolutional Neural Networks for Visual Recognition.” [Online]. Available: http://cs231n.github.io/convolutional-networks/. [Accessed: 12-Mar-2018].

[52] “Fundamentals of Deep Learning - Activation Functions and their use.” [Online]. Available: https://www.analyticsvidhya.com/blog/2017/10/fundamentals-deep-learning-activation-functions-when-to-use-them/. [Accessed: 29-Apr-2018].

[53] “Like that red skirt? Check this one! - Exponea Blog.” [Online]. Available: https://blog.exponea.com/like-that-red-skirt-check-this-one/. [Accessed: 12-Mar-2018].

[54] “Layers in convolutional neural network.” [Online]. Available: https://killianlevacher.github.io/blog/posts/post-2016-03-01/img/layeredRepresentation.jpg. [Accessed: 12-Mar-2018].

[55] “Max Pooling.” [Online]. Available: https://qph.fs.quoracdn.net/main-qimg-8afedfb2f82f279781bfefa269bc6a90. [Accessed: 29-Apr-2018].

[56] D. Ciresan, A. Giusti, L. Gambardella, and J. Schmidhuber, “Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images,” *Nips*, pp. 1–9, 2012.

[57] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9351, pp. 234–241, 2015.

[58] T. Naito, Y. Nagashima, K. Taira, N. Uchio, S. Tsuji, and J. Shimizu, “Identification and segmentation of myelinated nerve fibers in a cross-sectional optical microscopic image using a deep learning model,” *J. Neurosci. Methods*, vol. 291, pp. 141–149, Nov. 2017.

[59] R. Mesbah, B. McCane, and S. Mills, “Deep convolutional encoder-decoder for myelin and axon segmentation,” in *2016 International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 2016, pp. 1–6.

[60] Y. Y. Wang, Y. N. Sun, C. C. K. Lin, and M. S. Ju, “Segmentation of nerve fibers using multi-level gradient watershed and fuzzy systems,” *Artif. Intell. Med.*, vol. 54, no. 3, pp. 189–200, 2012.

[61] J. M. Ritchie, “On the relation between fibre diameter and conduction velocity in myelinated nerve fibres.,” *Proc. R. Soc. London. Ser. B, Biol. Sci.*, vol. 217, no. 1206, pp. 29–35, Dec. 1982.

[62] A. Brun and E. Englund, “A white matter disorder in dementia of the Alzheimer type: A pathoanatomical study,” *Ann. Neurol.*, vol. 19, no. 3, pp. 253–262, Mar. 1986.

[63] S. Ugrenović, I. Jovanović, L. Vasović, B. Kundalić, R. Čukuranović, and V. Stefanović, “Morphometric analysis of the diameter and g-ratio of the myelinated nerve fibers of the

human sciatic nerve during the aging process," *Anat. Sci. Int.*, vol. 91, no. 3, pp. 238–245, Jun. 2016.

[64] N. Stikov *et al.*, "In vivo histology of the myelin g-ratio with magnetic resonance imaging," *Neuroimage*, vol. 118, pp. 397–405, 2015.

[65] K. L. West, N. D. Kelm, R. P. Carson, and M. D. Does, "Quantitative analysis of mouse corpus callosum from electron microscopy images," *Data Br.*, vol. 5, pp. 124–128, Dec. 2015.

[66] I. O. Jelescu *et al.*, "In vivo quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy," *Neuroimage*, vol. 132, pp. 104–114, May 2016.

[67] K. L. West, N. D. Kelm, R. P. Carson, and M. D. Does, "A revised model for estimating g-ratio from MRI," *Neuroimage*, vol. 125, pp. 1155–1158, Jan. 2016.

[68] E. D. Schwartz, E. T. Cooper, C.-L. Chin, S. Wehrli, A. Tessler, and D. B. Hackney, "Ex vivo evaluation of ADC values within spinal cord white matter tracts.," *AJNR. Am. J. Neuroradiol.*, vol. 26, no. 2, pp. 390–7, Feb. 2005.

[69] N. Stikov *et al.*, "Quantitative analysis of the myelin g-ratio from electron microscopy images of the macaque corpus callosum," *Data Br.*, vol. 4, pp. 368–373, Sep. 2015.

[70] R. Gold, H.-P. Hartung, and K. V Toyka, "Animal models for autoimmune demyelinating disorders of the nervous system," *Mol. Med. Today*, vol. 6, no. 2, pp. 88–91, Feb. 2000.

[71] V. Schultz *et al.*, "Acutely damaged axons are remyelinated in multiple sclerosis and experimental models of demyelination," *Glia*, vol. 65, no. 8, pp. 1350–1360, Aug. 2017.

[72] L. Li, A. A. Velumian, M. Samoilova, and M. G. Fehlings, "A Novel Approach for Studying the Physiology and Pathophysiology of Myelinated and Non-Myelinated Axons in the CNS White Matter," *PLoS One*, vol. 11, no. 11, p. e0165637, Nov. 2016.

[73] M. F. Stidworthy, S. Genoud, U. Suter, N. Mantei, and R. J. M. Franklin, "Quantifying the Early Stages of Remyelination Following Cuprizone-induced Demyelination," *Brain Pathol.*, vol. 13, no. 3, pp. 329–339, Apr. 2006.

[74] R. C. Armstrong, A. J. Mierzwa, C. M. Marion, and G. M. Sullivan, "White matter

involvement after TBI: Clues to axon and myelin repair capacity," *Exp. Neurol.*, vol. 275, pp. 328–333, Jan. 2016.

[75] G. M. F. Costa, A. P. de Oliveira, P. M. Martinelli, E. R. da Silva Camargos, R. M. E. Arantes, and C. M. de Almeida-Leite, "Demyelination/remyelination and expression of interleukin-1β, substance P, nerve growth factor, and glial-derived neurotrophic factor during trigeminal neuropathic pain in rats," *Neurosci. Lett.*, vol. 612, pp. 210–218, Jan. 2016.

[76] K. Matsumuro, S. Izumo, F. Umehara, and M. Osame, "Chronic inflammatory demyelinating polyneuropathy: Histological and immunopathological studies on biopsied sural nerves," *J. Neurol. Sci.*, vol. 127, no. 2, pp. 170–178, Dec. 1994.

[77] K. Schmierer, F. Scaravilli, D. R. Altmann, G. J. Barker, and D. H. Miller, "Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain," *Ann. Neurol.*, vol. 56, no. 3, pp. 407–415, Sep. 2004.

[78] J. S. W. Campbell *et al.*, "Promise and pitfalls of g-ratio estimation with MRI," *Neuroimage*, 2017.

[79] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P. J. Basser, "Axcaliber: A method for measuring axon diameter distribution from diffusion MRI," *Magn. Reson. Med.*, vol. 59, no. 6, pp. 1347–1354, Jun. 2008.

[80] P. Kozlowski, D. Raj, J. Liu, C. Lam, A. C. Yung, and W. Tetzlaff, "Characterizing White Matter Damage in Rat Spinal Cord with Quantitative MRI and Histology," *J. Neurotrauma*, vol. 25, no. 6, pp. 653–676, Jun. 2008.

[81] J. P. Mottershead *et al.*, "High field MRI correlates of myelin content and axonal density in multiple sclerosis," *J. Neurol.*, vol. 250, no. 11, pp. 1293–1301, Nov. 2003.

[82] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *Cvpr 2015*, p. 5, 2015.

[83] J. U. Kim, H. G. Kim, and Y. M. Ro, "Iterative Deep Convolutional Encoder-Decoder Network for Medical Image Segmentation," Aug. 2017.

[84] S. Dodge and L. Karam, "Understanding how image quality affects deep neural networks,"

in *2016 8th International Conference on Quality of Multimedia Experience, QoMEX 2016*, 2016.

[85] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[86] L. Taylor and G. Nitschke, "Improving Deep Learning using Generic Data Augmentation," Aug. 2017.

[87] K. S. Sim, V. Teh, and M. E. Nia, "Adaptive noise Wiener filter for scanning electron microscope imaging system," *Scanning*, vol. 38, no. 2, pp. 148–163, Mar. 2016.

[88] H. S. Kushwaha, S. Tanwar, K. S. Rathore, and S. Srivastava, "De-noising Filters for TEM (Transmission Electron Microscopy) Image of Nanomaterials," in *2012 Second International Conference on Advanced Computing & Communication Technologies*, 2012, pp. 276–281.

[89] T. Tasdizen, R. Whitaker, R. Marc, and B. Jones, "Automatic Correction of Non-uniform Illumination in Transmission Electron Microscopy Images," 2005.

[90] J. Jo and Y. Bengio, "Measuring the tendency of CNNs to Learn Surface Statistical Regularities," Nov. 2017.

[91] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," Feb. 2015.

[92] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, 1993.

[93] N. Stikov *et al.*, "In vivo histology of the myelin g-ratio with magnetic resonance imaging," *Neuroimage*, vol. 118, pp. 397–405, Sep. 2015.

[94] I. Goodfellow *et al.*, "Generative Adversarial Nets," *Adv. Neural Inf. Process. Syst. 27*, pp. 2672–2680, 2014.

[95] N. Souly, C. Spampinato, and M. Shah, "Semi Supervised Semantic Segmentation Using Generative Adversarial Network."

[96] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, “Semantic Segmentation using Adversarial Networks.”

[97] S. Sankaranarayanan, Y. Balaji, A. Jain, S. N. Lim, and R. Chellappa, “Unsupervised Domain Adaptation for Semantic Segmentation with GANs,” Nov. 2017.